Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/006473

International filing date: 01 March 2005 (01.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/647,013
Filing date: 27 January 2005 (27.01.2005)

Date of receipt at the International Bureau: 20 May 2005 (20.05.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

1319233

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

May 11, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/647,013*
FILING DATE: *January 27, 2005*
RELATED PCT APPLICATION NUMBER: *PCT/US05/06473*

Certified by

Jon W. Dudas

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(c).

Docket Number 8940.6173		Type a plus sign (+) inside this box →	+

INVENTOR(s)/APPLICANT(s)

LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
Halenbeck Bosch Linnemann Lee	Robert Elizabeth Thomas Ernestine		San Rafael, CA Cupertino, CA San Francisco, CA Kensington, CA

TITLE OF INVENTION (500 characters max)

LEADER SEQUENCES FOR SECRETED POLYPEPTIDES AND METHODS FOR PRODUCTION THEREOF

CORRESPONDENCE ADDRESS

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, L.L.P., **Customer Number 22,852**

ENCLOSED APPLICATION PARTS (check all that apply)

☒ Specification: 105 Pages ☒ Drawings: 10 Sheets/9 Figures

☐ Other: [Number] Pages; [Description]

METHOD OF PAYMENT

☒	A check or money order is enclosed to cover the Provisional filing fees	PROVISIONAL FILING FEE $200.00 Total Number of Pages of specification, drawings,sequence or computer listing, or other papers 105. If more than 100 pages, add $250 for each additional 50 pages or fraction thereof. $450 (Size Fee) Reduction by ½ For Small Entity $[Number] TOTAL FILING FEE $450
☐	The Commissioner is hereby authorized to charge any deficiency or credit any overpayment i n the filing fees to Deposit Account Number 06-0916.	

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are:

Respectfully submitted,

SIGNATURE _____ Date January 27, 2005

TYPED OR PRINTED NAME Lisa M. Matovcik, Ph.D. REGISTRATION NO. 53,283

☐ Additional inventors are being named on separately numbered sheets attached hereto.

PROVISIONAL APPLICATION FILING ONLY

ATTORNEY DOCKET NO: 8940.6173
Customer No. 22,852

UNITED STATES PROVISIONAL PATENT APPLICATION

for

LEADER SEQUENCES FOR SECRETED POLYPEPTIDES AND METHODS FOR PRODUCTION THEREOF

by

ROBERT HALENBECK

ELIZABETH BOSCH

THOMAS LINNEMANN

and

ERNESTINE LEE

LEADER SEQUENCES FOR SECRETED POLYPEPTIDES AND METHODS FOR PRODUCTION THEREOF

FIELD OF THE INVENTION

[0001] The present invention relates to leader sequences that are useful for production of heterologous secreted polypeptides, nucleic acid constructs that encode such leader sequences and heterologous secreted polynucleotides, vectors that contain such nucleic acid constructs, recombinant host cells that contain such nucleic acid constructs, vectors and polypeptides, and methods of making and using such secreted polypeptides with such heterologous leader sequences.

BACKGROUND OF THE INVENTION

[0002] Proteins are the most prominent biomolecules in living organisms. In addition to their role as structural components and catalysts, they play a crucial role in regulatory processes. Both regulation of cell proliferation and metabolic functions are largely controlled and effected by the cooperation of numerous cellular and extracellular proteins. For example, signal transduction pathways of many kinds that affect critical physiological responses operate through proteins by way of their intermolecular interactions.

[0003] The extracellular proteins, sometimes referred to as the "secreted proteins," are likely to function as intercellular communicators of signals acting as ligands while their counterpart membrane associated receptors having extracellular and intracellular or cytoplasmic domains, transmit an extracellular signals into the cell upon ligand/receptor binding on the cell surfaces. Secreted proteins are typically expressed as full-length polypeptides, sometimes referred to as protein precursors, that are processed in the Golgi or the ER in the post-translational phase by cleavage of the secretory leader sequences to generate a mature polypeptide or by addition of carbohydrates in a glycosylation process (Hirschberg (1987)).

[0004] While receptors have been considered as important potential therapeutic targets, secreted proteins are of particular interest as potential therapeutic agents. Secreted proteins often have a signaling or hormone function, and hence have a high and specific biological activity (Schoen, F. J., (1994)). For example, secreted proteins control physiological reactions such as differentiation and proliferation, blood clotting and

thrombolysis, somatic growth and cell death, and immune response (Schoen, F. J., (1994)). Significant resources and research efforts have been expended for the discovery and investigation of new secreted proteins controlling biological functions. Many of such secreted proteins, including cytokines and peptide hormones, are manufactured and used as therapeutic agents (Zavyalov et al., (1997)). However, of the several thousand expected secreted proteins, few are currently used as therapeutic compounds

[0005] Secreted proteins are characterized by having a hydrophobic amino acid sequence at each of their N-terminus, a sequence that is generally referred to as a signal peptide (SP) or a secretory leader sequence although there are some secreted proteins such as the fibroblast growth factor family that lack the characteristic hydrophobic sequence. This SP is typically about 16 to 30 amino acid residues in length and is usually cleaved by a signal peptidase in the Golgi or the ER lumen before it is exported outside the cell. The resulting mature protein or the actual secreted polypeptide, thus, lacks the signal peptide sequence.

[0006] Naturally occurring secreted proteins are typically expressed in varying amounts depending on their physiological roles *in vivo*. As a result, many proteins, when expressed under the regulation of their naturally occurring secretory leader sequences are expressed in quantities that are too low for commercial purposes. It would be highly desirable, therefore, to be able to produce proteins for therapeutic applications in large quantities, regardless of how it is produced in the natural environment. It would, hence, be advantageous if nucleic acid constructs and methods could be devised to enable increased protein production *in vivo* or *in vitro*.

SUMMARY OF THE INVENTION

[0007] It is one of the objects of the present invention to provide nucleic acid and polypeptide constructs for producing proteins in higher yields than when such proteins are produced in their natural environment.

[0008] It is another one of the objects of the present invention to provide vectors, host cells and methods for producing proteins in higher yields than when such proteins are produced in their natural environment.

[0009] In accordance with one or more of the objects of the present invention, there is provided polypeptide or polynucleotide constructs as above where the

polypeptides and polynucleotides are modified, such as by formation of a fusion molecule using a fusion partner. The fusion molecules of the invention may be prepared by any conventional technique.

[0010] In accordance to one of the objectives, therefore, there is provided the present invention as embodied in the following examples:

[0011] 1. A heterologous polypeptide comprising a secretory leader and a mature polypeptide, wherein the secretory leader is operably linked to an N-terminus of the mature polypeptide, wherein the secretory leader is not so linked to the mature polypeptide in nature, and wherein the secretory leader comprises a leader sequence of a secreted protein, and the secreted protein is selected from Table 1.

[0012] 2. The heterologous polypeptide of 1, wherein the secreted protein is collagen type IX alpha 1 chain, long form or SEQ ID NO: 2.

[0013] 3. The heterologous polypeptide of 1, wherein the secreted protein is alpha-2-antiplasmin precursor (alpha-2-plasmin inhibitor) or SEQ ID NO: 3.

[0014] 4. The heterologous polypeptide of 1, wherein the secreted protein is trinucleotide repeat containing 5 or SEQ ID NO: 9.

[0015] 5. The heterologous polypeptide of 1, wherein the secreted protein is ARMET protein or SEQ ID NO: 19.

[0016] 6. The heterologous polypeptide of 1, wherein the secreted protein is calumenin or SEQ ID NO: 22.

[0017] 7. The heterologous polypeptide of 1, wherein the secreted protein is COL9A1 or SEQ ID NO: 26.

[0018] 8. The heterologous polypeptide of 1, wherein the secreted protein is NBL1 or SEQ ID NO: 28.

[0019] 9. The heterologous polypeptide of 1, wherein the secreted protein is PACAP protein or SEQ ID NO: 31.

[0020] 10. The heterologous polypeptide of 1, wherein the secreted protein is alpha-1B-glycoprotein precursor (alpha-1-B glycoprotein) or SEQ ID NO: 37.

[0021] 11. The heterologous polypeptide of 1, wherein the secreted protein is brain-specific angiogenesis inhibitor 2 precursor or SEQ ID NO: 41.

[0022] 12. The heterologous polypeptide of 1, wherein the secreted protein is SPOCK2 or SEQ ID NO: 47.

[0023] 13. The heterologous polypeptide of 1, wherein the secreted protein is protein disulfide-isomerase (EC 5341) ER60 precursor or SEQ ID NO: 54.

[0024] 14. The heterologous polypeptide of 1, wherein the secreted protein is serine or cysteine proteinase inhibitor, clade A (alpha-1) or SEQ ID NO: 57.

[0025] 15. The heterologous polypeptide of 1, wherein the secreted protein is GM2 ganglioside activator precursor or SEQ ID NO: 62.

[0026] 16. The heterologous polypeptide of 1, wherein the secreted protein is coagulation factor X precursor or SEQ ID NO: 69.

[0027] 17. The heterologous polypeptide of 1, wherein the secreted protein is secreted phosphoprotein 1 (osteopontin, bone sialoprotein 1) or SEQ ID NO: 75.

[0028] 18. The heterologous polypeptide of 1, wherein the secreted protein is Vitamin D-binding protein precursor or SEQ ID NO: 79.

[0029] 19. The heterologous polypeptide of 1, wherein the secreted protein is interleukin 6 (interferon, beta 2) or SEQ ID NO: 82.

[0030] 20. The heterologous polypeptide of 1, wherein the secreted protein is orosomucoid 1 precursor or SEQ ID NO: 86.

[0031] 21. The heterologous polypeptide of 1, wherein the secreted protein is hemopexin or SEQ ID NO: 88.

[0032] 22. The heterologous polypeptide of 1, wherein the secreted protein is glycoprotein hormones, alpha polypeptide precursor or SEQ ID NO: 94.

[0033] 23. The heterologous polypeptide of 1, wherein the secreted protein is kininogen 1 or SEQ ID NO: 97.

[0034] 24. The heterologous polypeptide of 1, wherein the secreted protein is prolyl 4-hydroxylase, beta subunit or SEQ ID NO: 102.

[0035] 25. The heterologous polypeptide of 1, wherein the secreted protein is proopiomelanocortin or SEQ ID NO: 104.

[0036] 26. The heterologous polypeptide of 1, wherein the secreted protein is prostaglandin D2 synthase or SEQ ID NO: 107.

[0037] 27. The heterologous polypeptide of 1, wherein the secreted protein is alpha-2-glycoprotein 1, zinc or SEQ ID NO: 111.

[0038] 28. The heterologous polypeptide of 1, wherein the secreted protein is chromogranin A or SEQ ID NO: 116.

[0039] 29. The heterologous polypeptide of 1, wherein the secreted protein is cystatin M precursor or SEQ ID NO: 120.

[0040] 30. The heterologous polypeptide of 1, wherein the secreted protein is clusterin isoform 1 or SEQ ID NO: 127.

[0041] 31. The heterologous polypeptide of 1, wherein the secreted protein is inter-alpha (globulin) inhibitor H1 or SEQ ID NO: 131.

[0042] 32. The heterologous polypeptide of 1, wherein the secreted protein is leukemia inhibitory factor (cholinergic differentiation factor) or SEQ ID NO: 137.

[0043] 33. The heterologous polypeptide of 1, wherein the secreted protein is lumican or SEQ ID NO: 140.

[0044] 34. The heterologous polypeptide of 1, wherein the secreted protein is secretoglobin, family 2A, member 2 or SEQ ID NO: 145.

[0045] 35. The heterologous polypeptide of 1, wherein the secreted protein is nov precursor or SEQ ID NO: 147.

[0046] 36. The heterologous polypeptide of 1, wherein the secreted protein is reticulocalbin 1 precursor SEQ ID NO: 153.

[0047] 37. The heterologous polypeptide of 1, wherein the secreted protein is reticulocalbin 2, EF-hand calcium binding domain or SEQ ID NO: 159.

[0048] 38. The heterologous polypeptide of 1, wherein the secreted protein is gastric intrinsic factor or SEQ ID NO: 167.

[0049] 39. The heterologous polypeptide of 1, wherein the secreted protein is cerberus 1 or SEQ ID NO: 175.

[0050] 40. The heterologous polypeptide of 1, wherein the secreted protein is lipocalin 2 (oncogene 24p3) or SEQ ID NO: 177.

[0051] 41. The heterologous polypeptide of 1, wherein the secreted protein is interleukin 18 binding protein isoform C precursor or SEQ ID NO: 181.

[0052] 42. The heterologous polypeptide of 1, wherein the secreted protein is cell growth regulator with EF hand domain 1 or SEQ ID NO: 185.

[0053] 43. The heterologous polypeptide of 1, wherein the secreted protein is leukocyte immunoglobulin-like receptor, subfamily A or SEQ ID NO: 189.

[0054] 44. The heterologous polypeptide of 1, wherein the secreted protein is spondin 2, extracellular matrix protein or SEQ ID NO: 191.

[0055] 45. The heterologous polypeptide of 1, wherein the secreted protein is transmembrane protein 4 or SEQ ID NO: 196.

[0056] 46. The heterologous polypeptide of 1, wherein the secreted protein is sparc/osteonectin, cwcv and kazal-like domain proteoglycan or SEQ ID NO: 200.

[0057] 47. The heterologous polypeptide of 1, wherein the secreted protein is Rho GTPase activating protein 25 isoform b or SEQ ID NO: 207.

[0058] 48. The heterologous polypeptide of 1, wherein the secreted protein is dickkopf homolog 3 or SEQ ID NO: 209.

[0059] 49. The heterologous polypeptide of 1, wherein the secreted protein is ameloblastin precursor or SEQ ID NO: 215.

[0060] 50. The heterologous polypeptide of 1, wherein the secreted protein is chorionic gonadotropin, beta polypeptide 8 precursor or SEQ ID NO: 218.

[0061] 51. The heterologous polypeptide of 1, wherein the secreted protein is multiple coagulation factor deficiency 2 or SEQ ID NO: 222.

[0062] 52. The heterologous polypeptide of 1, wherein the secreted protein is similar to common salivary protein 1 or SEQ ID NO: 227.

[0063] 53. The heterologous polypeptide of 1, wherein the secreted protein is hypothetical protein FLJ32115 or SEQ ID NO: 232.

[0064] 54. The heterologous polypeptide of 1, wherein the secreted protein is oncoprotein-induced transcript 3 or SEQ ID NO: 235.

[0065] 55. The heterologous polypeptide of 1, wherein the secreted protein is MGC40499 or SEQ ID NO: 239.

[0066] 56. The heterologous polypeptide of 1, wherein the secreted protein is interleukin 18 binding protein isoform A precursor or SEQ ID NO: 241.

[0067] 57. The heterologous polypeptide of 1, wherein the secreted protein is interleukin 1 receptor antagonist isoform 1 precursor or SEQ ID NO: 245.

[0068] 58. The heterologous polypeptide of 1, wherein the secreted protein is WFIKKN2 protein or SEQ ID NO: 248.

[0069] 59. The heterologous polypeptide of 1, where in the secreted protein is similar to hypothetical protein 9330140G23 or SEQ ID NO: 254.

[0070] 60. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 20 – 21.

[0071] 61. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 23 – 25.

[0072] 62. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence of SEQ ID NO: 27.

[0073] 63. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 32 – 36.

[0074] 64. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 38 – 40.

[0075] 65. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 48 – 53.

[0076] 66. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 76 – 78.

[0077] 67. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 80 – 81.

[0078] 68. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 83 – 85.

[0079] 69. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence of SEQ ID NO: 87.

[0080] 70. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 95 – 96.

[0081] 71. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence of SEQ ID NO: 103.

[0082] 72. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 108 – 110.

[0083] 73. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 112 – 115.

[0084] 74. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 117 – 119.

[0085] 75. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 121 – 126.

[0086] 76. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 128 – 130.

[0087] 77. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 132 – 136.

[0088] 78. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 138 – 139.

[0089] 79. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 141 – 144.

[0090] 80. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 154 – 158.

[0091] 81. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 160 – 166.

[0092] 82. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 178 – 180.

[0093] 83. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 186 – 188.

[0094] 84. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 197 – 199.

[0095] 85. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 210 – 214.

[0096] 86. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs:223 – 226.

[0097] 87. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 233 – 234.

[0098] 88. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence of SEQ ID NO: 240.

[0099] 89. The heterologous polypeptide of 1, wherein the leader sequence comprises an amino acid sequence selected from among SEQ ID NOs: 246 – 247.

[00100] 90. The heterologous polypeptide of 1, wherein the mature polypeptide is a secreted polypeptide, an extracellular portion of a transmembrane protein, or a soluble receptor.

[00101] 91. The heterologous polypeptide of 90, wherein the secreted polypeptide is a growth factor, a cytokine, a lymphokine, an interferon, a hormone, a stimulatory factor, an inhibitory factor, a soluble receptor or splice variants thereof.

[00102] 92. A secretory leader comprising a leader amino acid sequence chosen from among the leader sequences of Table 1 and Table 2.

[00103] 93. The secretory leader sequence of 92, wherein the leader amino acid sequence is chosen from Appendix A.

[00104] 94. The secretory leader sequence of 92, wherein the leader amino acid sequence comprises amino acid residues MKTCWKIPVFFFVCSFLEPWASA (SEQ ID NO: 1).

[00105] 95. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 4 – 8.

[00106] 96. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 10 – 18.

[00107] 97. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 20 – 21.

[00108] 98. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 23 – 25.

[00109] 99. The secretory leader of 92, wherein the leader amino acid sequence is SEQ ID NO: 27.

[00110] 100. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 29 – 30.

[00111] 101. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 32 – 36.

[00112] 102. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 38 – 40.

[00113] 103. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 42 – 46.

[00114] 104. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 48 – 53.

[00115] 105. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 55 – 56.

[00116] 106. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 58 – 61.

[00117] 107. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 63 – 67.

[00118] 108. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 69 – 74.

[00119] 109. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 76 – 78.

[00120] 110. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 80 – 81.

[00121] 111. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 83 – 85.

[00122] 112. The secretory leader of 92, wherein the leader amino acid sequence is SEQ ID NO: 87.

[00123] 113. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 89 – 93.

[00124] 114. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 95 – 96.

[00125] 115. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 98 – 101.

[00126] 116. The secretory leader of 92, wherein the leader amino acid sequence is SEQ ID NO: 103.

[00127] 117. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 105 – 106.

[00128] 118. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 109 – 110.

[00129] 119. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 112 – 115.

[00130] 120. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 117 – 119.

[00131] 121. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 121 – 126.

[00132] 122. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 128 – 130.

[00133] 123. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 132 – 136.

[00134] 124. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 138 – 139.

[00135] 125. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 141 – 144.

[00136] 126. The secretory leader of 92, wherein the leader amino acid sequence is SEQ ID NO: 146.

[00137] 127. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 148 – 152.

[00138] 128. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 154 – 158.

[00139] 129. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 160 -166.

[00140] 130. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 168 – 174.

[00141] 131. The secretory leader of 92, wherein the leader amino acid sequence is SEQ ID NO: 176.

[00142] 132. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 178 – 180.

[00143] 133. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 182 – 184.

[00144] 134. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 186 – 188.

[00145] 135. The secretory leader of 92, wherein the leader amino acid sequence is SEQ ID NO: 190.

[00146] 136. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 192 – 195.

[00147] 137. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 197 – 199.

[00148] 138. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 201 – 206.

[00149] 139. The secretory leader of 92, wherein the leader amino acid sequence is SEQ ID NO: 208.

[00150] 140. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 210 – 214.

[00151] 141. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 216 – 217.

[00152] 142. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 219 – 221.

[00153] 143. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 223 – 226.

[00154] 144. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 228 – 231.

[00155] 145. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 233- 234.

[00156] 146. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 236 – 238.

[00157] 147. The secretory leader of 92, wherein the leader amino acid sequence is SEQ ID NO: 240.

[00158] 148. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 242 – 244.

[00159] 149. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 246 – 247.

[00160] 150. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 249 – 253.

[00161] 151. The secretory leader of 92, wherein the leader amino acid sequence is chosen from among SEQ ID NOs: 255 – 256.

[00162] 152. The heterologous polypeptide of 1, further comprising a fusion partner.

[00163] 153. The heterologous polypeptide of 152, wherein the fusion partner is a polymer.

[00164] 154. The heterologous polypeptide of 153, wherein the polymer is a second polypeptide is selected from the group consisting of all or part of human serum albumin, fetuin A, fetuin B, and Fc.

[00165] 155. The heterologous polypeptide of 154, wherein the polymer is polyethylene glycol.

[00166] 156. A nucleic acid molecule comprising a polynucleotide that comprises a nucleotide sequence encoding the heterologous polypeptide of any one of 1 – 91 and 152 – 153 or the secretory leader of any one of 92 – 151.

[00167] 157. A nucleic acid molecule encoding a heterologous polypeptide, comprising a first polynucleotide that encodes the secretory leader of 92, a second polynucleotide that encodes a mature polypeptide, wherein the first polynucleotide and the second polynucleotide are operably linked to facilitate secretion of the heterologous polypeptide from a cell, and wherein the first and second polynucleotide are not so linked in nature.

[00168] 158. The nucleic acid of 157, wherein the mature polypeptide is a secreted polypeptide, an extracellular portion of a transmembrane protein, or a soluble receptor.

[00169] 159. The nucleic acid molecule of 157, further comprising a third polynucleotide, wherein the third polynucleotide is a Kozak sequence, GCCGCCACC, that is situated at its 5' end.

[00170] 160. The nucleic acid molecule of 157, further comprising a fourth polynucleotide, wherein the fourth polynucleotide comprises a restriction enzyme-cleavable sequence at its 3' end.

[00171] 161. The nucleic acid molecule of 157, further comprising a fifth polynucleotide that encodes a tag.

[00172] 162. The nucleic acid molecule of 161, wherein the tag is a purification tag.

[00173] 163. The nucleic acid molecule of 161, wherein the tag comprises at least one selected from V5, HisX6, HisX8, an avidin molecule, and a biotin molecule.

[00174] 164. The nucleic acid molecule of 161, further comprising a sixth polynucleotide that encodes a second cleavable sequence that can be cleaved by a second enzyme, wherein the second cleavable sequence is situated upstream of the tag if the tag is situated at the C-terminus of the heterologous polypeptide, or downstream of the tag if the tag is situated at the N terminus of the heterologous polypeptide.

[00175] 165. The nucleic acid molecule of 164, wherein the second enzyme is thrombin or TEV from a tobacco virus.

[00176] 166. A vector comprising the nucleic acid molecule of 156 or 157, further comprising an origin of replication and a selectable marker.

[00177] 167. The vector of 166, wherein the origin of replication is selected from the group consisting of SV40 ori, Pol ori, EBNA ori, and pMB1 ori.

[00178] 168. The vector of 166, wherein the selectable marker is an antibiotic resistance gene.

[00179] 169. The vector of 166, wherein the antibiotic resistance is selected from the group consisting of puromycin resistance, kanamycin resistance, and ampicillin resistance.

[00180] 170. A recombinant host cell comprising a cell and the heterologous polypeptide of any of 1 – 91, the nucleic acid molecule of any of 156-165, or the vector of any of 166 – 169.

[00181] 171. The recombinant host cell of 170, wherein the cell is a eukaryotic cell.

[00182] 172. The recombinant host cell of 170, wherein the cell is a human cell.

[00183] 173. A method of producing a secreted polypeptide, comprising:

[00184] (a) providing the nucleic acid molecule of any of 156-165; and

(b) expressing the nucleic acid molecule in an expression system.

[00185] 174. The method of 173, wherein the expression system is a cellular expression system or a cell free expression system.

[00186] 175. The method of 174, wherein the expression system is a cellular expression system and the cell is a mammalian cell.

[00187] 176. The method of 175, wherein the mammalian cell is a cell of a 293 cell line or a CHO cell line.

[00188] 177. The method of 176, wherein the 293 cell is a 293-T cell or a 293-6E cell.

Description of the Figures

[00189] **FIG.** 1 is an alignment of the amino acid sequences of: (a) a leader sequence of the present invention ("collagen_leader"); (b) a cDNA clone previously designated as MGC:21955 having an annotation of an unknown protein, and designated herein as CLN00517648; and (c) a publicly accessible sequence NP_001842_NM_001851, corresponding to collagen type IX alpha I chain, long form (*Homo Sapiens*). These sequences all start with methionine ("M") as amino acid residue 1 at the N terminus. This clone CLN00517648_5pv1 was sequenced and found to contain 253 amino acid residues.

[00190] **FIG.** 2 is a Western blot showing expression of the polypeptide in the conditioned medium of cultured 293E cells transfected with the cDNA of clone CLN00517648. The amount of protein expressed was compared to three (3) standards of V5-Hisx6 tagged Delta-like protein 1 extracellular protein and V5-Hisx6 tagged CSF-1

Receptor extracellular domain, shown in the three right hand lanes at 8, 33, and 133 nanograms/milliliter (ng/ml), respectively.

[00191] **FIG.** 3 is a diagrammatic representation of a starting vector pTT5 (4398 bps) kindly provided by Dr. Yves Durocher (Durocher, Y., *et al.* 2002).

[00192] **FIG.** 4 shows the sequence of Vector A for insertion into pTT5 vector to replace the region "ccdb" on pTT5. Vector A includes from left to right: a restriction site EcoR1, the gene of interest, "------" representing the open reading frame encoding a mature polypeptide of interest to be inserted, another restriction site, BamH1, a cleavable sequence represented by a sequence encoding thrombin, a tag represented by V5H8, and a random sequence with a stop codon.

[00193] **FIG.** 5 shows sequences for Vector B and Vector C to be inserted into pTT5 to replace the region "ccdb." Vector B includes, from left to right: a Kozak sequence, a leader sequence ("SP") such as the collagen leader sequence of the present invention, a EcoR1 site, "-----" representing the open reading frame of a mature polypeptide of interest, to be inserted, a BamH1 site, a tag such as V5H8, and a random sequence including a stop codon. Vector C includes, from left to right: a Kozak sequence, a leader sequence ("SP") such as the collagen leader sequence of the present invention, a EcoR1 site, "------" representing the open reading frame of a mature polypeptide of interest, to be inserted, a BamH1 site, a cleavable sequence represented by a sequence encoding thrombin, a tag such as V5H8, and a random sequence including a stop codon.

[00194] **FIG.** 6 shows sequences for Vector D and Vector E respectively. Vector D includes, from left to right: a restriction site EcoR1, "------ " representing the open reading frame encoding a mature polypeptide of interest to be inserted, another restriction site, BamH1, an Fc domain sequence followed by a stop codon. Vector E includes, from left to right: a Kozak sequence ("GCCGCCACC"), ATG of the secreted protein of interest representing the open reading frame encoding a mature polypeptide of interest to be inserted (less the ATG), a restriction site EcoR1, "------" representing the open reading frame of a mature polypeptide of interest, to be inserted, another restriction site, BamH1, an Fc domain sequence followed by a stop codon.

[00195] **FIG.** 7 is an example of a vector for making a stable puromycin resistant cell line. Specifically, the pTT2p vector includes, *inter alia*, murine polyoma signals to make an episomal pTT2-gateway vector.

[00196] **FIG.** 8 shows an SDS-PAGE analysis of protein expression, in CHO soy medium, employing 28 of the leader sequences described herein. The top two (2) panels show SDS-PAGE developed with coomassie stain and the bottom two (2) panels show SDS-PAGE developed with silver stain. Table 4, columns 6-11, identifies the specific leader sequence represented in each SDS-PAGE lane. As shown, a bovine serum albumin (BSA) standard was used at 8, 16, and 32 milligrams/liter (mg/L).

[00197] **FIG.** 9 shows an SDS-PAGE analysis of protein expression, in CHO soy medium, employing an additional 28 of the leader sequences described herein. The top two (2) panels show SDS-PAGE developed with coomassie stain and the bottom two (2) panels show SDS-PAGE developed with silver stain. Table 4, columns 6-11, identifies the specific leader sequence represented in each SDS-PAGE lane. As shown, a bovine serum albumin (BSA) standard was used at 8, 16, and 32 milligrams/liter (mg/L).

[00198] Table 1 lists information regarding the leader sequences employed in the invention. Column 1 shows an internal designation identification number, column 2 shows the reference identification number, column 3 shows the identified secreted protein, and column 4 shows an internal parameter "the treevote."

[00199] Table 2 shows lists information regarding the leader sequences employed in the invention. Column 1 shows an internal designation identification number, column 2 shows the SEQ ID NO. for each leader sequence (P1), column 3 shows the reference identification number, column 4 shows the type of leader sequence, i.e., full length and alternative leader sequences, and column 5 shows the identified secreted protein.

[00200] Table 3 shows a summary of results for tested leader sequences. Column 1 shows a clone designation identification number, column 2 shows the micrograms/milliliter (μg/ml) of protein detected in SDS-PAGE developed by coomassie stain, column 3 shows a rank of highest to lowest expression results for the leader sequences tested, column 4 shows a yes or no vote for whether a band was detected using SDS-PAGE developed by silver stain, column 5 shows the molecular weight of the tested leader sequences in Daltons, column 6 shows the gel number and lane number that

corresponds to Figures 8-9, column 7 shows an internal notation for the tested secreted proteins, column 8 shows a protein identification number, column 9 shows an internal designation identification number, column 10 shows a reference identification number, column 11 shows the identified secreted protein.

[00201] Appendix A shows the amino acid sequences of the leader sequences shown in Table 2 (i.e., P1 sequences).

Detailed Description of the Invention

[00202] As described herein, Applicants have observed that in order for some secreted proteins to express and secrete in larger quantities, a secretory leader sequence from another, i.e., different, secreted protein is desirable. Employing heterologous secretory leader sequences is advantageous in that a resulting mature amino acid sequence, i.e., protein, of the secreted polypeptide is not altered as the secretory leader sequence is removed in the endoplasmic reticulum (ER) during the secretion process. Moreover, the addition of a heterologous secretory leader is often required to express and secrete, for example, extracellular domains of Type II single transmembrane proteins (STM), as the secretory leader, which is also the transmembrane spanning domain, has to be removed to make them soluble.

[00203] Thus, to identify potential robust secretory leader sequence(s) that could universally be used for secreted proteins and to express the intracellular domain of Type II STMs, Applicants' have cloned and expressed, as described herein, many different secreted proteins and measured their expression and secretion levels in the supernatant of 293 mammalian cells (see, for example, Example 1, Figures 8-9, and Table 3). Several high expressors and high secretors proteins were observed.

[00204] In one embodiment, Applicants have identified a secretory leader sequence belonging to secreted protein collagen type IX alpha I chain, long form and selected this particular leader sequence to further examine its ability to promote expression and secretion when used as a heterologous secretory leader sequence. As described herein, the amino acid sequence of the secreted protein Collagen type IX alpha I chain, long form is predicted to be MKTCWKIPVFFFVCSFLEPWASA (SEQ ID NO: 2). As further described herein, vectors were constructed containing this particular secretory leader, several proteins were cloned removing the secretory leader from the full

length encoding sequence, and by cloning them into vectors containing SEQ ID NO: 2 resulting in secreted proteins with a heterologous secretory leader sequence. As further shown and described here, high expression and secretion of several selected proteins were also observed.

[00205] The present invention may be more clearly understood in light of the following definitions. Generally, the terms used herein have their ordinary meaning and the meanings given them specifically below.

[00206] The terms "polynucleotide," "nucleotide," "nucleic acid," "nucleotide molecule," "nucleic acid molecule," "nucleic acid sequence," "polynucleotide sequence," and "nucleotide sequence" are used interchangeably herein to refer to polymeric forms of nucleotides of any length. The polynucleotides can contain deoxyribonucleotides, ribonucleotides, and/or their analogs or derivatives. For example, nucleic acids can be naturally occurring DNA or RNA, or can be synthetic analogs, as known in the art. The terms also encompass genomic DNA, genes, gene fragments, exons, introns, regulatory sequences or regulatory elements (such as promoters, enhancers, initiation and termination regions, other control regions, expression regulatory factors, and expression controls), isolated DNA, and cDNA. The terms also encompass mRNA, tRNA, rRNA, ribozymes, splice variants, antisense RNA, antisense conjugates, RNAi, siRNA and isolated RNAs. The terms also encompass recombinant polynucleotides, heterologous polynucleotides, branched polynucleotides, labeled polynucleotides, hybrid DNA/RNA, polynucleotide constructs, vectors comprising the subject nucleic acids, nucleic acid probes, primers, and primer pairs. The polynucleotides can comprise modified nucleic acid molecules, with alterations in the backbone, sugars, or heterocyclic bases, such as methylated nucleic acid molecules, peptide nucleic acids, and nucleic acid molecule analogs, which may be suitable as, for example, probes if they demonstrate superior stability and/or binding affinity under assay conditions. Analogs of purines and pyrimidines, including radiolabeled and fluorescent analogs, are known in the art. The polynucleotides can have any three-dimensional structure. The terms also encompass single-stranded, double-stranded and triple helical molecules that are DNA, RNA, or hybrid DNA/RNA and that may encode a full-length gene or a biologically active fragment thereof. Biologically active fragments of polynucleotides can encode the

polypeptides herein, as well as anti-sense, ribozymes, or RNAi molecules. Thus, the full length polynucleotides herein may be treated with enzymes, such as Dicer, to generate a library of short RNAi fragments which are within the scope of the present invention.

[00207] The terms "polypeptide," "peptide," and "protein," used interchangeably herein, refer to a polymeric form of amino acids of any length, which can include naturally-occurring amino acids, coded and non-coded amino acids, chemically or biochemically modified, derivatized, or designer amino acids, amino acid analogs, peptidomimetics, and depsipeptides, and polypeptides having modified, cyclic, bicyclic, depsicyclic, or depsibicyclic peptide backbones. The term also includes conjugated proteins, fusion proteins, including, but not limited to, GST fusion proteins, fusion proteins with a heterologous amino acid sequence, fusion proteins with heterologous and homologous leader sequences, fusion proteins with or without N-terminal methionine residues, pegylated proteins, and immunologically tagged proteins. Also included in this term are variations of naturally occurring proteins, where such variations are homologous or substantially similar to the naturally occurring protein, as well as corresponding homologs from different species. Variants of polypeptide sequences include insertions, additions, deletions, or substitutions compared with the subject polypeptides. The term also includes peptide aptamers.

[00208] A "secretory leader," "signal peptide," or a "leader sequence," contain a sequence of amino acid residues, typically positioned at the N terminus of a polypeptide, which directs the intracellular trafficking of the polypeptide. Polypeptides that contain a secretory leader, signal peptide or leader sequence, typically also contain a secretory leader, signal peptide or leader sequence cleavage site. Such polypeptides, after cleavage at the cleavage sites, generate mature polypeptides, for example, after extracellular secretion or after being directed to the appropriate intracellular compartment.

[00209] A "secreted" protein refers to those proteins capable of being directed to the endoplasmic reticulum (ER), secretory vesicles, or the extracellular space as a result of a secretory leader, signal peptide or leader sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo

extracellular processing to produce a "mature" polypeptide. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

[00210] A "biologically active" entity, or an entity having "biological activity," is one having structural, regulatory, or biochemical functions of a naturally occurring molecule or any function related to or associated with a metabolic or physiological process. Biologically active polynucleotide fragments are those exhibiting activity similar, but not necessarily identical, to an activity of a polynucleotide of the present invention. The biological activity can include an improved desired activity, or a decreased undesirable activity. For example, an entity demonstrates biological activity when it participates in a molecular interaction with another molecule, such as hybridization, or when it has therapeutic value in alleviating a disease condition, or when it has prophylactic value in inducing an immune response to the molecule, or when it has diagnostic value in determining the presence of the molecule, such as a biologically active fragment of a polynucleotide that can be detected as unique for the polynucleotide molecule, or that can be used as a primer in PCR.

[00211] As noted above, a "biologically active" entity, or an entity having "biological activity," is one having structural, regulatory, or biochemical functions of a naturally occurring molecule or any function related to or associated with a metabolic or physiological process. Biologically active polypeptide fragments are those exhibiting activity similar, but not necessarily identical, to an activity of a polypeptide of the present invention. The biological activity can include an improved desired activity, or a decreased undesirable activity. For example, an entity demonstrates biological activity when it participates in a molecular interaction with another molecule, or when it has therapeutic value in alleviating a disease condition, or when it has prophylactic value in inducing an immune response to the molecule, or when it has diagnostic value in determining the presence of the molecule. A biologically active polypeptide or fragment thereof includes one that can participate in a biological reaction, for example, one that can serve as an epitope or immunogen to stimulate an immune response, such as production of antibodies, or that can participate in signal transduction by binding to receptors, proteins, or nucleic acids, activating enzymes or substrates.

[00212] An "isolated" or "substantially isolated" polynucleotide, or a polynucleotide in "substantially pure form," in "substantially purified form," or as an "isolate," is one that is substantially free of the sequences with which it is associated in nature, or other nucleic acid sequences that do not include a sequence or fragment of the subject polynucleotides. By substantially free is meant that less than about 90%, less than about 80%, less than about 70%, less than about 60%, or less than about 50% of the composition is made up of materials other than the isolated polynucleotide. For example, where at least about 99% of the total macromolecules is the isolated polynucleotide, the polynucleotide is at least about 99% pure, and the composition comprises less than about 1% contaminant.

[00213] "Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their desired function. Thus, a given promoter operably linked to a coding sequence is capable of effecting the expression of the coding sequence when the proper transcription factors, etc., are present. The promoter need not be contiguous with the coding sequence, so long as it functions to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between the promoter sequence and the coding sequence, as can translated introns, and the promoter sequence can still be considered "operably linked" to the coding sequence.

[00214] "Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, viral, synthetic, or synthetic origin which, by virtue of its origin or manipulation is not associated with all or a portion of the polynucleotide with which it is associated in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide.

[00215] A "control element" refers to a polynucleotide sequence which aids in the expression of a coding sequence to which it is linked. The term includes promoters, transcription termination sequences, upstream regulatory domains, polyadenylation signals, and when appropriate, leader sequences and enhancers, which collectively provide for the transcription and translation of a coding sequence in a host cell.

[00216] A "promoter" as used herein is a DNA regulatory region capable of binding RNA polymerase in a mammalian cell and initiating transcription of a downstream (3' direction) coding sequence operably linked thereto. For purposes of the present invention, a promoter sequence includes the minimum number of bases or elements necessary to initiate transcription of a gene of interest at levels detectable above background. Within the promoter sequence is a transcription initiation site, as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eucaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes. Promoters further include those that are naturally contiguous to a nucleic acid molecule and those that are not naturally contiguous to a nucleic acid molecule. Additionally, a promoter includes inducible promoters, conditionally active promoters, such as a cre-lox promoter, constitutive promoters and tissue specific promoters.

[00217] By "selectable marker" is meant a gene which confers a phenotype on a cell expressing the marker, such that the cell can be identified under appropriate conditions. Generally, a selectable marker allows selection of transformed cells based on their ability to thrive in the presence or absence of a chemical or other agent that inhibits an essential cell function. Suitable markers, therefore, include genes coding for proteins which confer drug resistance or sensitivity thereto, impart color to, or change the antigenic characteristics of those cells transfected with a molecule encoding the selectable marker, when the cells are grown in an appropriate selective medium. For example, selectable markers include: cytotoxic markers and drug resistance markers, whereby cells are selected by their ability to grow on media containing one or more of the cytotoxins or drugs; auxotrophic markers by which cells are selected by their ability to grow on defined media with or without particular nutrients or supplements, such as thymidine and hypoxanthine; metabolic markers by which cells are selected for, e.g., their ability to grow on defined media containing the appropriate sugar as the sole carbon source, or markers which confer the ability of cells to form colored colonies on chromogenic substrates or cause cells to fluoresce.

[00218] "Transformation," as used herein, refers to the insertion of an exogenous polynucleotide into a host cell, irrespective of the method used for insertion: for example, transformation by direct uptake, transfection, infection, and the like. For

particular methods of transfection, see further below. The exogenous polynucleotide may be maintained as a nonintegrated vector, for example, an episome, or alternatively, may be integrated into the host genome.

[00219] A "gene," for the purposes of the present disclosure, includes a DNA region encoding a gene product, as well as all DNA regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. Accordingly, a gene includes, but is not necessarily limited to, promoter sequences, terminators, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites and locus control regions.

[00220] "Gene expression" refers to the conversion of the information, contained in a gene, into a gene product. A gene product can be the direct transcriptional product of a gene (*e.g.*, mRNA, tRNA, rRNA, antisense RNA, ribozyme, structural RNA or any other type of RNA) or a protein produced by translation of an mRNA. Gene products also include RNAs which are modified, by processes such as capping, polyadenylation, methylation, and editing, and proteins modified by, for example, methylation, acetylation, phosphorylation, ubiquitination, ADP-ribosylation, myristilation, and glycosylation.

[00221] A "coding sequence" or a sequence which "encodes" a selected polypeptide, is a nucleic acid molecule which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to, cDNA from viral, procaryotic or eucaryotic mRNA, genomic DNA sequences from viral (e.g. DNA viruses and retroviruses) or procaryotic DNA, and especially synthetic DNA sequences. A transcription termination sequence may be located 3' to the coding sequence.

[00222] By "fragment" is intended a polypeptide consisting of only a part of the intact full-length polypeptide sequence and structure. The fragment can include a C-terminal deletion an N-terminal deletion, and/or an internal deletion of the native polypeptide. A fragment of a protein will generally include at least about 5-10 contiguous amino acid residues of the full-length molecule, preferably at least about 15-

25 contiguous amino acid residues of the full-length molecule, and most preferably at least about 20-50 or more contiguous amino acid residues of the full-length molecule, or any integer between 5 amino acids and the full-length sequence.

[00223] The term "binds specifically," in the context of antibody binding, refers to high avidity and/or high affinity binding of an antibody to a specific polypeptide, or more accurately, to an epitope of a specific polypeptide. Antibody binding to such epitope on a polypeptide can be stronger than binding of the same antibody to any other epitopes, particularly other epitopes that can be present in molecules in association with, or in the same sample as the polypeptide of interest. For example, when an antibody binds more strongly to one epitope than to another, adjusting the binding conditions can result in antibody binding almost exclusively to the specific epitope and not to any other epitopes on the same polypeptide, and not to any other polypeptide, which does not comprise the epitope. Antibodies that bind specifically to a subject polypeptide may be capable of binding other polypeptides at a weak, yet detectable, level (e.g., 10% or less of the binding shown to the polypeptide of interest). Such weak binding, or background binding, is readily discernible from the specific antibody binding to a subject polypeptide, e.g., by use of appropriate controls. In general, antibodies of the invention bind to a specific polypeptide with a binding affinity of 10^{-7} M or greater (e.g., 10^{-8} M, 10^{-9} M, 10^{-10}, 10^{-11}, etc.).

[00224] The term "host cell" or "recombinant host cell" includes an individual cell, cell line, cell culture, or a cell *in vivo*, which can be or has been a recipient of any polynucleotides or polypeptides of the invention, for example, a recombinant vector, an isolated polynucleotide, antibody or fusion protein. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology, physiology, or in total DNA, RNA, or polypeptide complement) to the original parent cell due to natural, accidental, or deliberate mutation and/or change. Host cells can be prokaryotic or eukaryotic, including mammalian, insect, amphibian, reptile, crustacean, avian, fish, plant and fungal cells. A host cell includes cells transformed, transfected, transduced, or infected *in vivo* or *in vitro* with a polynucleotide of the invention, for example, a recombinant vector. A host cell which comprises a recombinant vector of the invention may be called a "recombinant host cell."

[00225] The terms "modulator" or "agent" refers to a substance that binds to or modulates a level or activity of a subject polypeptide or a level of mRNA encoding a subject protein or DNA, or that modulates the activity of a cell containing the subject protein or nucleic acids. Where the agent modulates a level of mRNA encoding a subject protein, agents include ribozymes, antisense, and RNAi molecules, including siRNA. Where the agent is a substance that modulates a level of activity of a subject polypeptide, agents include antibodies specific for the subject polypeptide, peptide aptamers, small molecules, agents that bind a ligand-binding site in a subject polypeptide, and the like. Antibody agents include antibodies that interfere with or specifically bind to a subject polypeptide and activate the polypeptide, such as receptor-ligand binding that initiates signal transduction; antibodies that specifically bind a subject polypeptide and inhibit binding of another molecule to the polypeptide, thus preventing activation of a signal transduction pathway; antibodies that bind a subject polypeptide to modulate transcription; antibodies that bind a subject polypeptide to modulate translation; as well as antibodies that bind a subject polypeptide on the surface of a cell to initiate antibody-dependent cytotoxicity ("ADCC") or to initiate cell killing or cell growth. Small molecule agents include those that bind the polypeptide to modulate activity of the polypeptide or cell containing the polypeptide in a similar fashion. The term "agent" also refers to substances that modulate a condition or disorder associated with a subject polynucleotide or polypeptide. Such agents include subject polynucleotides themselves, subject polypeptides themselves, and the like. Agents may be chosen from amongst candidate agents, as defined below.

[00226] The terms "candidate modulator," "candidate agent," or "test agent," used interchangeably herein, encompass numerous chemical classes, typically synthetic, semi-synthetic, or naturally occurring inorganic or organic molecules, small molecules, macromolecular complexes or antibodies. Candidate agents can be small organic compounds having a molecular weight of more than about 50 and less than about 2,500 daltons. Candidate agents can comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and can include at least an amine, carbonyl, hydroxyl or carboxyl group, and can contain at least two of the functional chemical groups. The candidate agents can comprise cyclical carbon or

heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules, including oligonucleotides, polynucleotides, and fragments thereof, depsipeptides, polypeptides and fragments thereof, oligosaccharides, polysaccharides and fragments thereof, lipids, fatty acids, steroids, purines, pyrimidines, derivatives thereof, structural analogs, modified nucleic acids, modified, derivatized or designer amino acids, or combinations thereof. An agent which modulates a biological activity of a subject polypeptide increases or decreases the activity at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 50%, at least about 100%, or at least about 2-fold, at least about 5-fold, or at least about 10-fold or more when compared to a suitable control.

[00227] The term "agonist" refers to a substance that mimics the function of an active molecule. Agonists include, but are not limited to, drugs, hormones, antibodies, and neurotransmitters, as well as analogues and fragments thereof.

[00228] The term "antagonist" refers to a molecule that competes for the binding sites of an agonist, but does not induce an active response. Antagonists include, but are not limited to, drugs, hormones, antibodies, and neurotransmitters, as well as analogues and fragments thereof.

[00229] The term "receptor" refers to a polypeptide that binds to a specific extracellular molecule and may initiate a cellular response.

[00230] The term "ligand" refers to any molecule that binds to a specific site on another molecule.

[00231] An agent that "modulates the level of expression of a nucleic acid" in a cell is one that brings about an increase or decrease of at least about 1.25-fold, at least about 1.5-fold, at least about 2-fold, at least about 5-fold, at least about 10-fold, or more in the level (i.e., an amount) of mRNA and/or polypeptide following cell contact with a candidate agent compared to a control lacking the agent.

[00232] An "antibody" herein refers to an immunoglobulin molecule or an active fragment of such, including for example, a Fab fragment, a variable or constant region of a heavy chain, a variable or constant region of a light chain, a complementarity determining region (cdr), or a framework region. Thus, the antibody can be a

monoclonal antibody, a polyclonal antibody, or a single chain antibody. The antibody can also be a neutralizing antibody, an agonist, or an antagonist. The antibody can be a fusion molecule linked to a cytotoxic molecule. The antibody can comprise a TCR, a fibronectin, a CTLA4Ig or other backbone.

[00233] A "humanized" antibody is an antibody that contains mostly human immunoglobulin sequences. This term is generally used to refer to a non-human immunoglobulin that has been modified to incorporate portions of human sequences. A humanized antibody may include a human antibody that contains entirely human immunoglobulin sequences.

[00234] A "composition" of modulators, polypeptides, or polynucleotides herein refers to a composition that usually contains a pharmaceutically acceptable carrier or excipient that is conventional in the art and which is suitable for administration into a subject for therapeutic, diagnostic, or prophylactic purposes. For example, compositions for oral administration can form solutions, suspensions, tablets, pills, capsules, sustained release formulations, oral rinses, or powders.

[00235] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed. Moreover, it must be understood that the invention is not limited to the particular embodiments described, as such may, of course, vary. Further, the terminology used to describe particular embodiments is not intended to be limiting, since the scope of the present invention will be limited only by its claim.

[00236] Unless defined otherwise, the meanings of all technical and scientific terms used herein are those commonly understood by one of ordinary skill in the art to which this invention belongs. One of ordinary skill in the art will also appreciate that any methods and materials similar or equivalent to those described herein can also be used to practice or test the invention. Further, all publications mentioned herein are incorporated by reference.

[00237] It must be noted that, as used herein and in the appended claims, the singular forms "a," "or," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a subject polypeptide" includes a

plurality of such polypeptides and reference to "the agent" includes reference to one or more agents and equivalents thereof known to those skilled in the art, and so forth.

[00238] Further, all numbers expressing quantities of ingredients, reaction conditions, % purity, polypeptide and polynucleotide lengths, and so forth, used in the specification and claims, are modified by the term "about," unless otherwise indicated. Accordingly, the numerical parameters set forth in the specification and claims are approximations that may vary depending upon the desired properties of the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits, applying ordinary rounding techniques.

[00239] Nonetheless, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value, however, inherently contains certain errors from the standard deviation of its experimental measurement.

[00240] All publications cited are incorporated by reference herein in their entireties, including references cited in such publications are also incorporated by reference in their entireties.

Leader Sequences

[00241] As described herein, Applicants have identified secretory leader sequences from secreted proteins useful for producing proteins in higher yields than when such protein are produced in their natural environment. Identified secretory leader sequences, described herein include, for example, interleukin-9 precursor, T cell growth factor P40, P40 cytokine, triacylglycerol lipase, pancreatic precursor, somatoliberin precursor, vasopressin-neurophysin 2-copeptin precursor, beta-enoendorphin-dynorphin precursor, complement C2 precursor, small inducible cytokine A14 precursor, elastase 2A precursor, plasma serine protease inhibitor precursor, granulocyte-macrophage colony-stimulating factor precursor, interleukin-2 precursor, interleukin-3 precursor, alpha-fetoprotein precursor, alpha-2-HS-glycoprotein precursor, serum albumin precursor, inter-alpha-trypsin inhibitor light chain, serum amyloid P-component precursor, apolipoprotein A-II precursor, apolipoprotein D precursor, colipase precursor, carboxypeptidase A1 precursor, alpha-s1 casein precursor, beta casein precursor, cystatin

SA precursor, follitropin beta chain precursor, glucagon precursor, complement factor H precursor, histidine-rich glycoprotein precursor, interleukin-5 precursor, alpha-lactalbumin precursor, Von Ebner's gland protein precursor, matrix Gla-protein precursor, alpha-1-acid glycoprotein 2 precursor, phospholipase A2 precursor, dendritic cell chemokine 1, statherin precursor, transthyretin precursor, apolipoprotein A-1 precursor, apolipoprotein C-III precursor, apolipoprotein E precursor, complement component C8 gamma chain precursor, serotransferrin precursor, beta-2-microglobulin precursor, neutrophils defensins 1 precursor, triacylglycerol lipase gastric precursor, haptoglobin precursor, neutrophils defensins 3 precursor, neuroblastoma suppressor of tumorigenicity 1 precursor, small inducible cytokine A13 precursor, CD5 antigen-like precursor, phospholipids transfer protein precursor, dickkopf related protein-4 precursor, elastase 2B precursor, alpha-1-acid glycoprotein 1 precursor, beta-2-glycoprotein 1 precursor, neutrophils gelatinase-associated lipocalin precursor, C-reactive protein precursor, interferon gamma precursor, kappa casein precursor, plasma retinol-binding protein precursor, interleukin-13 precursor, and any of the secreted proteins set forth in Tables 1-3.

[00242] The above-identified secretory leader sequences, vectors and methods described herein, are useful in the expression of a wide variety of polypeptides, including, for example, secreted polypeptides, extracellular proteins, transmembrane proteins, and receptors, such as a soluble receptor. Examples of such polypeptides include cytokines and growth factors, such as Interleukins 1 through 18, the interferons, the lymphokines, hormones, RANTES, lymphotoxin-β, Fas ligand, flt-3 ligand, ligand for receptor activator of NF-kappa B (RANKL), soluble receptors, TNF-related apoptosis-inducing ligand (TRAIL), CD40 ligand, Ox40 ligand, 4-1BB ligand (and other members of the TNF family), thymic stroma-derived lymphopoietin, stimulatory factors, such as, for example, granulocyte colony stimulating factor and granulocyte-macrophage colony stimulating factor, inhibitory factors, mast cell growth factor, stem cell growth factor, epidermal growth factor, growth hormone, tumor necrosis factor, leukemia inhibitory factor, oncostatin-M, splice variants, and hematopoietic factors such as erythropoietin and thrombopoietin.

[00243] Descriptions of some proteins that can be expressed according to the invention may be found in, for example, Human Cytokines: Handbook for Basic and Clinical Research, Vol. II (Aggarwal and Gutterman, eds. Blackwell Sciences, Cambridge Mass., 1998); Growth Factors:A Practical Approach (McKay and Leigh, eds., Oxford University Press Inc., New York, 1993) and The Cytokine Handbook (A W Thompson, ed.; Academic Press, San Diego Calif.; 1991).

[00244] Receptors for any of the aforementioned proteins may also be expressed using secretory leader sequences, vectors and methods described herein, including, for example, both forms of tumor necrosis factor receptor (referred to as p55 and p75), Interleukin-1 receptors (type 1 and 2), Interleukin-4 receptor, Interleukin-15 receptor, Interleukin-17 receptor, Interleukin-18 receptor, granulocyte-macrophage colony stimulating factor receptor, granulocyte colony stimulating factor receptor, receptors for oncostatin-M and leukemia inhibitory factor, receptor activator of NF-kappa B (RANK), receptors for TRAIL, and receptors that comprise death domains, such as Fas or Apoptosis-Inducing Receptor (AIR).

[00245] Other proteins that can be expressed using the secretory leader sequences, vectors and methods described herein include, for example, cluster of differentiation antigens (referred to as CD proteins), for example, those disclosed in Leukocyte Typing VI (Proceedings of the VIth International Workshop and Conference; Kishimoto, Kikutani et al., eds.; Kobe, Japan, 1996), or CD molecules disclosed in subsequent workshops. Examples of such molecules include CD27, CD30, CD39, CD40; and ligands thereto (CD27 ligand, CD30 ligand and CD40 ligand). Several of these are members of the TNF receptor family, which also includes 41BB and OX40; the ligands are often members of the TNF family (as are 4-1BB ligand and OX40 ligand); accordingly, members of the TNF and TNFR families can also be expressed using the present invention.

[00246] Proteins that are enzymatically active may also be expressed employing the herein described secretory leader sequences, vectors and methods and include, for example, metalloproteinase-disintegrin family members, various kinases (including streptokinase and tissue plasminogen activator as well as Death Associated Kinase Containing Ankyrin Repeats, and IKR 1 and 2), TNF-alpha Converting Enzyme, and

numerous other enzymes. Ligands for enzymatically active proteins can also be expressed by applying the instant invention.

[00247] The secretory leader sequences, vectors and methods described herein, are also useful for the expression of other types of recombinant proteins, including, for example, immunoglobulin molecules or portions thereof, and chimeric antibodies (i.e., an antibody having a human constant region couples to a murine antigen binding region) or fragments thereof. Numerous techniques are known by which DNA encoding immunoglobulin molecules can be manipulated to yield DNAs capable of encoding recombinant proteins such as single chain antibodies, antibodies with enhanced affinity, or other antibody-based polypeptides (see, for example, Larrick et al., Biotechnology 7:934-938, 1989; Reichmann et al., Nature 332:323-327, 1988; Roberts et al., Nature 328:731-734, 1987; Verhoeyen et al., Science 239:1534-1536, 1988; Chaudhary et al., Nature 339:394-397, 1989).

Vectors, Host Cells, and Protein Production

[00248] The present invention provides recombinant vectors that contain, for example, nucleic acid constructs that encode secretory leader sequences and a selected heterologous polypeptide of interest, and host cells that are genetically engineered with the recombinant vectors. Selected heterologous polypeptides of interest in the present invention include, for example, an extracellular fragment of a secreted protein, a type I membrane protein, a type II membrane protein, a multi-membrane protein, and a soluble receptor. These vectors and host cells can be used for the production of polypeptides described herein, including fragments thereof by conventional recombinant techniques. The vector may be, for example, a phage, plasmid, viral or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

[00249] The polynucleotides may be joined to a vector containing a secretory leader sequence (see, for example, Table 1), and a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged *in vitro* using an appropriate packaging cell line and then transduced into host cells.

[00250] The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

[00251] As indicated, the expression vectors will typically include at least one selectable marker. Such markers include dihydrofolate reductase, G418, neomycin, or puromycin resistance for eukaryotic cell culture and tetracycline, kanamycin, puromycin, or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293 (including 293-6E and 293-T) and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

[00252] Among vectors useful in the present invention are the herein described vectors employing a pTT vector backbone, see, for example, Figures 3-7 (Durocher et al. (2002)). Briefly, the pTT vector backbone may be prepared by obtaining pIRESpuro/EGFP (pEGFP) and pSEAP basic vector(s), for example from Clontech (Palo Alto, CA), and pcDNA3.1, pCDNA3.1/Myc-$(His)_6$ and pCEP4 vectors can be obtained from, for example, Invitrogen. SuperGlo GFP variant (sgGFP) can be obtained from Q-Biogene (Carlsbad, CA). Preparing a pCEP5 vector can be accomplished by removing the CMV promoter and polyadenylation signal of pCEP4 by sequential digestion and self-ligation using *Sal*I and *Xba*I enzymes resulting in plasmid pCEP4Δ. A GblII fragment from pAdCMV5 (Massie et al., (1998)), encoding the CMV5-poly(A) expression cassette ligated in *Bgl*II-linearized pCEP4Δ, resulting in pCEP5 vector. The pTT vector can be prepared by deleting the hygromycin (*Bsm*I and *Sal*I excision followed

by fill-in and ligation) and EBNA1 (*Cla*I and *Nsi*I excision followed by fill-in and ligation) expression cassettes. The ColEI origin (*Fsp*I-*Sal*I fragment, including the 3' end of β-lactamase ORF) can be replaced with a *FspI-SalI* fragment from pcDNA3.1 containing the pMBI oring (and the same 3' end of β-lactamase ORF). A Myc-$(His)_6$ C-terminal fusion tag can be added to SEAP (*Hind*III-*Hpa*I fragment from pSEAP-basic) following in-frame ligation in pcDNA3.1/Myc-His digested with *Hind*III and *Eco*RV. Plasmids can subsequently be amplified in Escherichia coli (DH5α) grown in LB medium and purified using MAXI prep columns (Qiagen, Mississauga, Ontario, Canada). To quantify, plasmids can be subsequently diluted in 50 mM Tris-HCl pH 7.4 and absorbencies can be measured at 260nm and 280nm. Preferably, plasmid preparations with A_{260}/A_{280} ratios between about 1.75 and about 2.00 are used.

[00253] Introduction of a construct into a host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). After transfection of the vector or DNA construct encoding the present polypeptides into host cells, the cells can be allowed to grow to produce the present polypeptides.

[00254] A variety of host-expression vector systems may be utilized to express polypeptides of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express a polypeptide of the invention. These include, but are not limited to, microorganisms such as bacteria (e.g., E. coli, B. subtilis) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing polypeptide coding sequences; yeast (e.g., Saccharomyces, Pichia) transformed with recombinant yeast expression vectors containing polypeptide coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing polypeptide coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g.,

Ti plasmid) containing polypeptide coding sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 2936E, 293T, and 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

[00255] Typically, a heterologous polypeptide, whether modified or unmodified, may be expressed, as described above, or as a fusion protein, and may include not only secretion signals, but preferably also a secretory leader sequence (Table 1). A secretory leader sequence of the invention, directs certain proteins to the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

[00256] Proteins targeted to the ER by a secretory leader sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space--a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

[00257] Additionally, peptide moieties and/or purification tags may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others are familiar and routine techniques in the art. Suitable purification tags include, for example, V5, HISX6, HISX8, avidin, and biotin. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to stabilize and purify proteins. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins containing various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part

in a fusion protein is thoroughly advantageous for use in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified in the advantageous manner described. This is the case when Fc portion proves to be a hindrance to use in therapy and diagnosis, for example when the fusion protein is to be used as antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. See, Bennett et al., *J. Molecular Recognition*, 8:52-58 (1995) and Johanson et al, *J. Biol. Chem.*, 270:9459-9471 (1995).

[00258] A heterologous polypeptide of the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and preferably mammalian cells, or a cell free expression system. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Modifications

[00259] The invention encompasses polypeptides which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, $NABH_4$; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin.

[00260] Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of procaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

[00261] Also provided by the invention are chemically modified derivatives of the polypeptides of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Pat. No. 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

[00262] A polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired

therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

[00263] The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., Exp. Hematol. 20:1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

[00264] One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

Fusion Molecules of the Invention

[00265] In a further embodiment of the invention, the heterologous polypeptides of the present invention may be combined with one or more fusion partners to form fusion molecules. Such fusion molecules may advantageously provide improved pharmacokinetic properties when compared to an unmodified non-fused molecule. Modified derivatives of a selected heterologous polypeptide may be prepared by one skilled in the art, given the disclosures herein. Suitable chemical moieties for derivatization of a heterologous polypeptide include, for example, polymers, such as water soluble polymers, all or part of human serum albumin, fetuin A, fetuin B, and an Fc region.

[00266] Polymers, and in particular water soluble polymers, are useful in the present invention as the polypeptide to which each polymer is attached will not precipitate in an aqueous environment, such as a physiological environment. Preferably, polymers employed in the invention will be pharmaceutically acceptable for the preparation of a therapeutic product or composition. One skilled in the art will be able to select the desired polymer based on such considerations as whether the polymer/protein conjugate will be used therapeutically and, if so, the desired dosage, circulation time and resistance to proteolysis.

[00267] Suitable, clinically acceptable, water soluble polymers include, but are not limited to, polyethylene glycol (PEG), polyethylene glycol propionaldehyde, copolymers of ethylene glycol/propylene glycol, monomethoxy-polyethylene glycol, carboxymethylcellulose, dextran, polyvinyl alcohol (PVA), polyvinyl pyrrolidone, poly-1,3-dioxolane, poly-1,3,6-trioxane, ethylene/maleic anhydride copolymer, poly (β-amino acids) (either homopolymers or random copolymers), poly(n-vinyl pyrrolidone) polyethylene glycol, polypropylene glycol homopolymers (PPG) and other polyakylene oxides, polypropylene oxide/ethylene oxide copolymers, polyoxyethylated polyols (POG) (e.g., glycerol) and other polyoxyethylated polyols, polyoxyethylated sorbitol, or polyoxyethylated glucose, colonic acids or other carbohydrate polymers, Ficoll or dextran and mixtures thereof.

[00268] As used herein, polyethylene glycol (PEG) is meant to encompass any of the forms that have been used to derivatize other proteins, such as mono-(C1-C10)

alkoxy- or aryloxy-polyethylene glycol. Polyethylene glycol propionaldehyde may have advantages in manufacturing due to its stability in water.

[00269] Specifically, a modified heterologous polypeptide of the invention may be prepared by attaching polyaminoacids or branch point amino acids to the polypeptide. For example, the polyaminoacid may be a carrier protein that serves to increase the circulation half life of the polypeptide (i.e., in addition to the advantages achieved via a fusion molecule). For the therapeutic purpose of the present invention, such polyaminoacids should ideally be those that have or do not create neutralizing antigenic response, or other adverse responses. Such polyaminoacids may be selected from the group consisting of serum album (such as human serum albumin), an additional antibody or portion thereof, for example the Fc region, fetuin A, fetuin B, or other polyaminoacids, e.g. lysines. As described herein, the location of attachment of the polyaminoacid may be at the N-terminus, or C-terminus, or other places in between, and also may be connected by a chemical "linker" moiety to the selected molecule.

[00270] Polymers used herein, for example water soluble polymers, may be of any molecular weight and may be branched or unbranched. The polymers each typically have an average molecular weight of between about 2 kDa to about 100 kDa (the term "about" indicating that in preparations of a polymer, some molecules will weigh more, some less, than the stated molecular weight). The average molecular weight of each polymer is preferably is between about 5 kDa and about 50 kDa, more preferably between about 12 kDa and about 25 kDa. Generally, the higher the molecular weight or the more branches, the higher the polymer:protein ratio. Other sizes may be used, depending on the desired therapeutic profile, for example the duration of sustained release; the effects, if any, on biological activity; the ease in handling; the degree or lack of antigenicity and other known effects of a polymer on a modified molecule of the invention.

[00271] Polymers employed in the present invention are typically attached to a heterologous polypeptide with consideration of effects on functional or antigenic domains of the polypeptide. In general, chemical derivatization may be performed under any suitable condition used to react a protein with an activated polymer molecule. Activating groups which can be used to link the polymer to the active moieties include the

following: sulfone, maleimide, sulfhydryl, thiol, triflate, tresylate, azidirine, oxirane and 5-pyridyl.

[00272] Polymers of the invention are typically attached to a heterologous polypeptide at the alpha (α) or epsilon (ε) amino groups of amino acids or a reactive thiol group, but it is also contemplated that a polymer group could be attached to any reactive group of the protein that is sufficiently reactive to become attached to a polymer group under suitable reaction conditions. Thus, a polymer may be covalently bound to a heterologous polypeptide via a reactive group, such as a free amino or carboxyl group. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residue. Those having a free carboxyl group may include aspartic acid residues, glutamic acid residues and the C-terminal amino acid residue. Those having a reactive thiol group include cysteine residues.

[00273] Methods for preparing fusion molecules conjugated with polymers, such as water soluble polymers, will each generally contain the steps of: (a) reacting a heterologous polypeptide with a polymer under conditions whereby the polypeptide becomes attached to one or more polymers and (b) obtaining the reaction product. Reaction conditions for each conjugation may be selected from any of those known in the art or those subsequently developed, but should be selected to avoid or limit exposure to reaction conditions such as temperatures, solvents and pH levels that would inactivate the protein to be modified. In general, the optimal reaction conditions for the reactions will be determined case-by-case based on known parameters and the desired result. For example, the larger the ratio of polymer:polypeptide conjugate, the greater the percentage of conjugated product. The optimum ratio (in terms of efficiency of reaction in that there is no excess unreacted polypeptide or polymer) may be determined by factors such as the desired degree of derivatization (e.g., mono-, di-tri- etc.), the molecular weight of the polymer selected, whether the polymer is branched or unbranched and the reaction conditions used. The ratio of polymer (e.g., PEG) to a polypeptide will generally range from 1:1 to 100:1. One or more purified conjugates may be prepared from each mixture by standard purification techniques, including among others, dialysis, salting-out, ultrafiltration, ion-exchange chromatography, gel filtration chromatography and electrophoresis.

[00274] One may specifically desire an N-terminal chemically modified protein. One may select a polymer by molecular weight, branching, etc., the proportion of polymers to protein (polypeptide or peptide) molecules in the reaction mix, the type of reaction to be performed, and the method of obtaining the selected N-terminal chemically modified protein. The method of obtaining the N-terminal chemically modified protein preparation (i.e., separating this moiety from other monoderivatized moieties if necessary) may be by purification of the N-terminal chemically modified protein material from a population of chemically modified protein molecules.

[00275] Selective N-terminal chemical modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved. For example, one may selectively attach a polymer to the N-terminus of the protein by performing the reaction at a pH which allows one to take advantage of the pKa differences between the ε-amino group of the lysine residues and that of the α-amino group of the N-terminal residue of the protein. By such selective derivatization, attachment of a polymer to a protein is controlled: the conjugation with the polymer takes place predominantly at the N-terminus of the protein and no significant modification of other reactive groups, such as the lysine side chain amino groups, occurs. Using reductive alkylation, the polymer may be of the type described above and should have a single reactive aldehyde for coupling to the protein. Polyethylene glycol propionaldehyde, containing a single reactive aldehyde, may also be used.

[00276] In one embodiment, the present invention contemplates the chemically derivatized polypeptide to include mono- or poly- (e.g., 2-4) PEG moieties. "Pegylation" may be carried out by any of the pegylation reactions known in the art. Methods for preparing a pegylated protein product will generally include the steps of: (a) reacting a polypeptide with polyethylene glycol (such as a reactive ester or aldehyde derivative of PEG) under conditions whereby the protein becomes attached to one or more PEG groups; and (b) obtaining the reaction product(s). In general, the optimal

reaction conditions for the reactions will be determined case by case based on known parameters and the desired result.

[00277] There are a number of PEG attachment methods available to those skilled in the art. See, for example, EP 0 401 384; Malik et al., *Exp. Hematol.*, 20:1028-1035 (1992); Francis, *Focus on Growth Factors*, 3(2):4-10 (1992); EP 0 154 316; EP 0 401 384; WO 92/16221; WO 95/34326; and the other publications cited herein that relate to pegylation, the disclosures of which are hereby incorporated by reference.

[00278] The step of pegylation as described herein may be carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule. Thus, protein products according to the present invention include pegylated proteins wherein the PEG group(s) is (are) attached via acyl or alkyl groups. Such products may be mono-pegylated or poly-pegylated (e.g., containing 2-6, and preferably 2-5, PEG groups). The PEG groups are generally attached to the protein at the α- or ε-amino groups of amino acids, but it is also contemplated that the PEG groups could be attached to any amino group attached to the protein that is sufficiently reactive to become attached to a PEG group under suitable reaction conditions.

[00279] Pegylation by acylation generally involves reacting an active ester derivative of polyethylene glycol (PEG) with a polypeptide of the invention. For acylation reactions, the polymer(s) selected typically have a single reactive ester group. Any known or subsequently discovered reactive PEG molecule may be used to carry out the pegylation reaction. A preferred activated PEG ester is PEG esterified to N-hydroxysuccinimide (NHS). As used herein, "acylation" is contemplated to include, without limitation, the following types of linkages between the therapeutic protein and a polymer such as PEG: amide, carbamate, urethane, and the like, see for example, Chamow, *Bioconjugate Chem.*, 5 (2):133-140 (1994). Reaction conditions may be selected from any of those known in the pegylation art or those subsequently developed, but should avoid conditions such as temperature, solvent and pH that would inactivate the polypeptide to be modified.

[00280] Pegylation by acylation will generally result in a poly-pegylated protein. Preferably, the connecting linkage will be an amide. Also preferably, the resulting product will be substantially only (e.g., >95%) mono, di- or tri-pegylated. However,

some species with higher degrees of pegylation may be formed in amounts depending on the specific reaction conditions used. If desired, more purified pegylated species may be separated from the mixture (particularly unreacted species) by standard purification techniques, including among others, dialysis, salting-out, ultrafiltration, ion-exchange chromatography, gel filtration chromatography and electrophoresis.

[00281] Pegylation by alkylation generally involves reacting a terminal aldehyde derivative of PEG with a polypeptide in the presence of a reducing agent. For the reductive alkylation reaction, the polymer(s) selected should have a single reactive aldehyde group. An exemplary reactive PEG aldehyde is polyethylene glycol propionaldehyde, which is water stable, or mono C1-C10 alkoxy or aryloxy derivatives thereof, see for example, U.S. Pat. No. 5,252,714.

[00282] Additionally, heterologous polypeptides of the present invention and the epitope-bearing fragments thereof described herein can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These particular fusion molecules facilitate purification and show an increased half-life in vivo. This has been shown, e.g., for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins, for example, EP A 394,827; Traunecker et al., *Nature*, 331:84-86 (1988). Fusion molecules that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than, for example, a monomeric polypeptide or polypeptide fragment alone, see, for example, Fountoulakis et al., *J. Biochem.*, 270:3958-3964 (1995).

[00283] In another described embodiment, a human serum albumin fusion molecule may also be prepared as described herein and as further described in U.S. Patent No. 6,686,179.

[00284] Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide that facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, Calif., 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance,

hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984)).

[00285] Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

[00286] It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

Examples

Example 1: Expression of biologically active mature secreted proteins using a Cell-free system.

[00287] A nucleotide primer is designed and synthesized that contains the following nineteen nucleotides "5'CCACCCACCACCACCAATG 3'" followed by the first nineteen nucleotides predicted to encode the amino terminus of a mature secreted protein. To express the mature secreted protein, a second reverse primer is designed to a region of the plasmid approximately 1000 nucleotides downstream from the coding sequence of the gene to be expressed. The second primer is designed as the reverse complement of the vector sequence in this region such that this primer will be useful for doing PCR amplification of the mature coding sequence of the mature open reading frame to be expressed. The second primer is typically 17-23 nucleotides in length with a Tm of approximately 55-65°C.

[00288] A purified plasmid containing the cDNA to be expressed or E coli cells containing the plasmid that contains the cDNA to be expressed is then added as template to a standard PCR reaction that includes the two primers described above, standard PCR reagents, and a DNA polymerase that has proof-reading activity and subjected to 15-30 cycles of PCR amplification. The product of this PCR reaction is called "PCR1 coding template."

[00289] A separate PCR reaction is setup to prepare a "GST-Mega primer" that will be used to create a GST-fusion expression template. Using a plasmid template that contains the coding sequence for GST downstream of the Non-Omega translation

initiation sequence, a PCR reaction is prepared using the primer 5' GGTGACACTATAGAACTCACCTATCTCCCCAACA 3' and the primer 5' GGGCCCCTGGAACAGAACTTC 3' and amplified in a standard PCR reaction that includes the two primers described above, standard PCR reagents, and a DNA polymerase that has proof-reading activity and subjected to 15-30 cycles of PCR amplification. After the PCR reaction is complete the PCR product is subjected to exonuclease I treatment for 30 minutes at 37°C, then heat-inactivated at 80°C for 30 minutes, and the PCR product purified by agarose gel electrophoresis and extracted using a gel purification kit (Amersham) to produce the "GST-Mega primer."

[00290] The "GST-Mega primer" described above is then used to create GST-fusion expression template by combining it with the product of the first PCR reaction (PCR1 coding template) containing the mature coding of the cDNA to be expressed. An aliquot of the PCR1 coding template (0.5ul) is mixed with an aliquot of the GST-Mega primer (1ul) and a primer 5' GCGTAGCATTTAGGTGACACT 3' that encodes part of the SP6 promoter sequence and anneals to the five prime end of the GST Mega primer, and a second primer that is designed to a region of the plasmid approximately 300-350 nucleotides downstream from the coding sequence of the gene to be expressed. This second primer is designed as the reverse complement of the vector sequence in this region such that this primer will be useful for doing PCR amplification of the PCR1 coding template. This second primer is typically 17-23 nucleotides in length with a Tm of approximately 55-65°C. The "GST-fusion expression template" is then generated by doing a standard PCR reaction using standard PCR reagents, and a DNA polymerase that has proof-reading activity and subjected to 15-30 cycles of PCR amplification. The product of this PCR reaction is called "GST-fusion expression template."

[00291] An in vitro transcription reaction (50ul) is then prepared using 5ul of the GST-fusion expression template in the following buffer, 80 mM Hepes KOH pH 7.8, 16 mM Mg(OAc)2, 2 mM spermidine, 10 mM DTT containing 1 unit of SP6 (Promega) and 1 unit of RNasin (Promega) and incubated for 3 hours at 37°C. The mRNA is then subjected to ethanol precipitation by addition of 200ul of RNase-free water, 37.5 ul of 5M ammonium acetate, and 862ul of 99% ethanol, mixed by vortexing and then pelleted

by centrifugation at 15,000 x g for 10minutes at 4°C. The mRNA pellet is then washed in 70% ethanol and again pelleted by centrifugation at 15,000 x g for 5 minutes at 4°C.

[00292] For the in vitro translation reaction a stock of 2x Dialysis Buffer is prepared that contains 20 mM Hepes buffer pH 7.8 (KOH), 200 mM KOAc, 5.4 mM Mg(OAc)2, 0.8 mM Spermidine, 100 micomolar DTT, 2.4 mM ATP, 0.5 mM GTP, 32 mM creatine phosphate, 0.02 % NaN3, and 0.6 mM Amino Acid Mix minus ASP,TRP, GLU, ISO, LEU, PHE, and TYR. The amino acids ASP,TRP, GLU, ISO, LEU, PHE, and TYR are prepared separately as an 80 mM stock in 1N HCL and after complete dissolution are added to a final concentration of 0.6 mM. After addition of all ingredients the 2x Dialysis Buffer stock is adjusted to pH 7.6 using 5N KOH, filter sterilized, and stored frozen in aliquots at -80°C.

[00293] To resuspend the in vitro transcribed mRNA that has been ethanol precipitated and washed in 70% ethanol a 50ul translation mixture is prepared that includes Wheat Germ Reagent at a final OD 260nm of 60 plus the volume of 1x Dialysis Buffer (to which 2 mM DTT has been added) that brings the final volume to 50ul (Wheat Germ Reagent already includes 1x Dialysis Buffer in it). After removing the ethanol from the precipitated mRNA the 50ul translation mixture is added, allowed to sit for 5-10 minutes and then the mRNA is resuspended. The complete translation mixture containing the resuspended mRNA is then layered under 250ul of 1x Dialysis Buffer that had already been added to one well of 96 well round bottom microtiter plate to setup the Bilayer Reaction. The plate is then sealed manually with a plate seal and then incubated for 20 hours at 26°C.

[00294] To recover the recombinant protein expressed as a GST fusion, the translation mixture is transferred to a tube, diluted five-fold with phosphate buffer-saline containing 0.25 M sucrose, 2 mM DTT, and 10ul of glutathione-sepharose is added and incubated with mixing for 3 hours at 4°C. The sepharose beads containing the bound GST fusion protein are then washed three times in phosphate buffer-saline containing 0.25 M sucrose and 2 mM DTT. A fourth wash is then done in protease cleavage buffer containing 50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 2 mM DTT, and 0.25 M sucrose. After careful removal of the wash buffer 10ul of final wash buffer is added back plus 0.4 ul of Prescission Protease (Amersham), the beads gently suspended with a

pipette, and then allowed to incubate over night at 4°C. To recover the cleaved secrete protein product, 20ul of final wash buffer is added and entire liquid fraction recovered by pipette or by filtering through a scintered frit. To stabilize the recovered secreted protein, purified BSA prepared as a 10mg/ml stock in PBS is added to a final concentration of 1 mg/ml and the protein sample then dialyzed in PBS and filter sterilized for storage prior to testing for biological activity. To produce additional protein the single Bilayer Reaction can be reproduced many times and the purification and formulation scaled accordingly. Typically, sixteen Bilayer Reactions will produce sufficient biologically active protein for testing in many biological assays.

[00295] Example 2 Identification of secreted proteins secreted from mammalian cells at high levels.

[00296] cDNAs were predicted bioinformatically to encode secreted proteins based upon a defined set of attributes that included, for example, the presence of a signal peptide typically encoded by the first 6-27 amino acid codons (18-81 nucleotides) of the open reading frame (ORF), beginning with 1-4 polar amino acids followed by a stretch of hydrophobic amino acids and then a short region of charged amino acids just before the cleavage site. Using this criteria, in addition to other physical characteristics, the signal peptide sequence of an unknown protein was determined that defines the cDNA as encoding a secreted protein.

[00297] In order to identify signal peptide(s) that yield high level protein secretion, a set of cDNAs predicted to encode secreted proteins were subcloned into a pTT5 expression vector in frame with a C-terminal V5 and His x 8 epitope and transiently transfected into 293T cells using a 96-well high throughput system. Purified plasmid DNA for each clone was prepared using the Qiagen™ Turbo DNA system in 96 well plates. The DNA concentration for each clone was determined by absorbance at 260nm and diluted to50 ug/mL. For transient transfection of ten 96-well plates, 10µl of each DNA plasmid was combined with 50µl of GIBCO Opti-MEM I (Cat#:319-85-070) in a round bottom 96-well polystyrene plate (named the master transfection plate). In order to generate the transfection complex, 37.5µl of Opti-MEM I preincubated for 5 minutes with 2.5µl of Fugene 6 (Roche Applied Science cat#:1988387), was added and the complex was allowed to form at room temp for about 30 minutes.

[00298] The transfection complex was subsequently diluted by the addition of 100µl of Opti-MEM I, mixed several times by pipetting in an up and down motion, and then transferred 20µl at a time into ten 96 well flat bottom poly-lysine-coated plates (Becton Dickinson cat#: 356461). 293T cells suspension (200µl at 2 x 10E5 cell/mL) in DMEM medium containing 10% FBS and penicillin and streptomycin were then added to each well and incubated at 37°C in 5 % CO_2. After approximately 40 hours, the medium was removed by aspiration, the cells briefly washed with 150µl phosphate-buffered saline (PBS), and then new pre-warmed medium was added. For measuring the expression and secretion level of each protein fresh HyQ-PF CHO Liquid Soy (Hyclone Cat# SH30359.02) medium (150µl) added to each well incubated at 37°C in 5 % CO_2. For measuring activity of secreted protein fresh DMEM medium containing 5% FBS and penicillin and streptomycin (150µl) was added in place of the HyQ-PF CHO Liquid Soy.

[00299] After an additional 48 hours the culture supernatant from all ten 96-well plates were harvested and combined into a single sterile deep well plate, covered with a sterile lid and centrifuged at 1400 RPM for 10 minutes to pellet any loose cells or cell debris. The supernatant was then transferred to a new sterile deep well plate for testing for protein expression by Western blot. The remaining cell layer on the plates was solubilized with 0.2% SDS, 0.5% NP-40 in PBS.

[00300] The expression of cDNAs in 293-6E cells was tested either using the high throughput transfection process, described above, or in larger quantities using 293-6E cells grown in shake flasks. For the high throughput process 293-6E cells were treated in an identical fashion as 293T cells. For scale-up expression, 293-6E cells were grown in polycarbonate Erlenmeyer flasks fitted with a vented screw cap and rotated on a table top shaker at 100 RPM in Freestyle Medium (Invitrogen®) at 37°C in 5% CO_2 at cell densities ranging from 0.5 to 3 x 10^6. Typically 50ml of culture was grown in a 250 ml flask. One day prior to setting up a transfection, 293-6E cells were diluted into fresh Freestyle medium to 0.6 x 10^6 cells/ml. On the day of transfection the cells were predicted to be in log phase (0.8 – 1.5 x 10^6 cells/ml) and adjusted to 10^6 cells/ml.

[00301] To prepare the transfection mix, 2.5 ml sterile PBS was added to two 15 ml tubes, into one 50µg DNA was added, into the other 100µl PEI solution (1 mg/ml sterile stock solution, Polyethylenimine, linear, 25 kDa., pH 7.0 (from Polysciences,

Warrington, WI) was added, the solutions were then combined and allowed to incubate for 15 minutes at room temp to form the transfection complex. The transfection mixture was then transferred to 293-6E suspension culture and allowed to grow for 4 -6 days at 37°C in 5% CO_2.

[00302] To determine protein secretion levels culture supernatants were analyzed by Western blot. Samples (15µl) were resolved by SDS-PAGE on a 26 lane Criterion gel (BioRad) and transferred to nitrocellulose, blocked, and then probed with an anti-V5 HRP conjugate (Invitrogen®). Secretion levels were determined by comparing band intensity to that of one of three different purified standards run on the same Western analysis at three different concentrations. The standards used were either 1) V5-Hisx6 tagged Delta-like protein 1 extracellular protein or 2) V5-Hisx6 tagged CSF-1 Receptor extracellular domain, each expressed separately using the baculovirus expression system and purified to > 90% purity, or 3) Positope (Invitrogen, cat#: R900-50) containing a V5 Hisx6 tag, each run separately or combined.

[00303] From the analysis of the high throughput expression of many cDNAS in 293T cells, several cDNAS were identified that resulted in very high secretion levels. The signal peptide sequence from one of the high expressing clones, CLN00517648 that encoded human collagen, type IX, alpha 1, long form was used to engineer the high level secretion of low-expressing cDNAs, type I TM proteins, and type II cDNAs by replacing the endogenous signal peptide sequence of each cDNA with that of collagen type IX, alpha 1. Constructs encoding human CD30 Ligand, SCDFR1, Ox40 Ligand, were engineered in the pTT5 vector and transfected into 293T and 293-6E cells to test expression and secretion using the improved signal peptidein 293T cells and in 293-6E cells using both the high throughput and the scale-up procedures.

References

[00304] The specification is most thoroughly understood in light of the following references, all of which are hereby incorporated by reference in their entireties. The disclosures of the patents and other references cited above are also hereby incorporated by reference.

1. Agrawal, S., Crooke, S.T. eds. (1998) Antisense Research and Application (Handbook of Experimental Pharmacology, Vol 131). Springer-Verlag New York, Inc.
2. Andreeff, M., Pinkel, D. eds. (1999) Introduction to Fluorescence In Situ Hybridization: Principles and Clinical Applications. John Wiley & Sons.
3. Ansel, H.C., Allen, L., Popovich, N.G. eds. (1999) Pharmaceutical Dosage Forms and Drug Delivery Systems. 7th ed. Lippencott Williams and Wilkins Publishers.
4. Beigelman, L., Karpeisky, A., Matulic-Adamic, J., Haeberli, P., Sweedler, D., Usman, N. (1995) Synthesis of 2'-modified nucleotides and their incorporation into hammerhead ribozymes. *Nucleic Acids Res.* 23:4434-4442.
5. Chen, S.Y., Bagley, J., Marasco, W.A. (1994) Intracellular antibodies as a new class of therapeutic molecule for gene therapy. *Hum. Gene Ther.* 5:595-601.
6. Cheng, W.F., Hung, C.F., Chai, C.Y., Hsu, K.F., He, L., Ling, M., Wu, T.C. (2001) Tumor-specific immunity and angiogenesis generated by a DNA vaccine encoding calreticulin linked to a tumor antigen. *J. Clin. Invest.* 108:669-678.
7. Chien, C., Bartel, P.L., Sternglanz, R., Fields S. (1991) The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci.* 88:9578-9581.
8. Coligan, J.E. et al., eds. (2002) Current Protocols in Immunology, vols. 1-4, including quarterly suppl.) John Wiley and Sons, Inc.New York, New York.
9. Deutscher, M.P., Simon, M.I., Abelson, J.N., eds. (1990) Guide to Protein Purification: Methods in Enzymology. (Methods in Enzymology Series, Vol 182). Academic Press.
10. Dieffenbach, C.W., Dveksler, G.S., eds. (1995) PCR Primer: A Laboratory Manual. Cold Spring Harbor Laboratory Press.
11. Durocher, Y. et al. (2002) Nucleic Acids Res. 30(2): High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells.

12. Fields, S., Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* 340:245-246.

13. Fukuhara, A. et al. (2004) Visfatin: A protein secreted by visceral fat that mimics the effects of insulin. Sciencexpress @ www.sciencexpress.org/16 December 2004/Page 1/10.1126/science.1097243.

14. Furth, P.A., Shamay, A., Wall, R.J., Hennighausen, L. (1992) Gene transfer into somatic tissues by jet injection. *Anal. Biochem.* 205:365-368.

15. Gaudilliere, B., Shi, Y., Bonni, A. (2002) RNA interference reveals a requirement for MEF2A in activity-dependent neuronal survival. *J. Biol. Chem.* 277:46,442-46,446.

16. Gennaro, A., ed. (2000) Remington: The Science and Practice of Pharmacy. 20th ed. Lippincott, Williams, & Wilkins.

17. Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I., Howard, B.H. (1982) The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eucaryotic cells by DNA-mediated transfection. *Proc. Natl. Acad. Sci.* 79:6777-6781.

18. Grosschedl, R., Baltimore, D. (1985) Cell-type specificity of immunoglobulin gene expression is regulated by at least three DNA sequence elements. *Cell* 41:885-897.

19. Grosveld, F., Kollias, G., eds. (1992) Transgenic Animals. 1st ed. Academic Press.

20. Harlow, E., Lane, D., eds. (1988) Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory.

21. Harlow, E., Lane, D., Harlow, E., eds. (1998) Using Antibodies: A Laboratory Manual: Portable Protocol NO. I. Cold Spring Harbor Laboratory.

22. Hartmann, G., Endres, S., eds. (1999) Manual of Antisense Methodology (Perspectives in Antisense Science). 1st ed. Kluwer Law International.

23. Hassanzadeh, G.H.G., De Silva, K.S., Dambly-Chudiere, C., Brys, L., Ghysen, A., Hamers, R., Muyldermans, S., De Baetselier, P. (1998) Isolation and characterization of single-chain Fv genes encoding antibodies specific for Drosophila Poxn protein. *FEBS Lett.* 437:75-80.

24. Heiser, A., Coleman, D., Dannull, J., Yancey, D., Maurice, M.A., Lallas, C.D., Dahm, P., Niedzwiecki, D., Gilboa, E., Vieweg, J. (2002) Autologous dendritic cells transfected with prostate-specific antigen RNA stimulate CTL responses against metastatic prostate tumors. *J. Clin. Invest.* 109:409-417.
25. Hirschberg, C. (1987) Annu. Rev. Biochem. 56:63-87.
26. Hoogenboom, H.R., de Bruin, A.P., Hufton, S.E., Hoet, R.M., Arends, J.W., Roovers, R.C. (1998) Antibody phage display technology and its applications. *Immunotechnology* 4:1-20.
27. Howard, G.C., Bethell, D.R. (2000) Basic Methods in Antibody Production and Characterization. CRC Press.
28. Jameson, D.M., Sawyer, W.H. (1995) Fluorescence anisotropy applied to biomolecular interactions. *Methods Enzymol.* 246:283-300.
29. Jia, S.H. et al. (2004) Pre-B cell colony-enhancing factor inhibits neutrophils apoptosis in experimental inflammation and clinical sepsis. J. Clin. Investigation 113(9): 1318 – 1327.
30. Jones, P., ed. (1998a) Vectors: Cloning Applications: Essential Techniques, John Wiley & Son, Ltd.
31. Jones, P., ed. (1998b) Vectors: Expression Systems: Essential Techniques, John Wiley & Son, Ltd.
32. Jost, C.R., Kurucz I., Jacobus, C.M., Titus, J.A., George, A.J., Segal, D.M. (1994) Mammalian expression and secretion of functional single-chain Fv molecules. *J. Biol. Chem.* 269:26,267-26,273.
33. Kabat, E.A., Wu T.T. (1991) Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. *J. Immunol.* 147:1709-1719.
34. Kibbe, A.H., ed. (2000) Handbook of Pharmaceutical Excipients. 3rd ed. Pharmaceutical Press.
35. Kirkpatrick, K.L., Mokbel, K. (2001) The significance of human telomerase reverse transcriptase (hTERT) in cancer. *Eur. J. Surg. Oncol.* 27:754-760.

36. Knutson, K.L., Schiffman, K., Disis, M.L. (2001) Immunization with a HER-2/neu helper peptide vaccine generates HER-2/neu CD8 T-cell immunity in cancer patients. *J. Clin. Invest.* 107:477-484.

37. Kolonin, M.G., Finley, R.L. Jr. (1998) Targeting cyclin-dependent kinases in *Drosophila* with peptide aptamers. *Proc. Natl. Acad. Sci.* 95:14,266-14,271.

38. Liu A.Y., Robinson R.R., Hellstrom K.E., Murray E.D. Jr., Chang C.P., Hellstrom I. (1987a) Chimeric mouse-human IgG1 antibody that can mediate lysis of cancer cells. *Proc. Natl. Acad. Sci.* 84:3439-3443.

39. Liu, A.Y., Robinson, R.R., Murray, E.D. Jr., Ledbetter, J.A., Hellstrom, I., Hellstorm, K.E. (1987b) Production of a mouse-human chimeric monoclonal antibody to CD20 with potent Fc-dependent biologic activity. *J. Immunol.* 139:3521-3526.

40. Machiels, J.P., van Baren, N., Marchand, M. (2002) Peptide-based cancer vaccines. *Semin. Oncol.* 29:494-502.

41. Massie, B., et al., (1998) Inducible overexpression of a toxic protine by an adenovirus vector with a tetracycline-regulatable expression cassette, *J. Virol.*, 72, 2289-2296).

42. Matz, M.V., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zaraisky, A.G., Markelov, M.L., Lukyanov, S.A. (1999) Fluorescent proteins from nonbioluminescent *Anthozoa* species. *Nat. Biotechnol.* 17:969-973.

43. Mayer, B.J. (2001) SH3 domains: complexity in moderation. *J. Cell Sci.* 114:1253-1263.

44. Milligan, J.F., Matteucci, M.D., Martin, J.C. (1993) Current concepts in antisense drug design. *J. Med. Chem.* 36:1923-1937.

45. Mitchell, D.A., Nair, S.K. (2000) RNA-transfected dendritic cells in cancer immunotherapy. *J. Clin. Invest.* 106:1065-1069.

46. Mitsumoto, Y. et al. (1991) Differential expression of the GLUT1 and GLUT4 glucose transporters during differentiation of L6 muscle cells. *Biochem. Biophys. Res. Commun.* 175: 652-9.

47. Mitsumoto, Y. et al. (1992) Development regulation of the subcellular distribution and glycosylation of GLUT1 and GLUT4 glucose transporters during myogenesis of L6 muscle cells. *J. Biol. Chem.* 267: 4957-4962.

48. Okayama, H., Berg, P. (1983) A cDNA cloning vector that permits expression of cDNA inserts in mammalian cells. *Mol. Cell. Biol.* 3:280-289.

49. O'Neil, N.J., Martin, R.L., Tomlinson, M.L., Jones, M.R., Coulson, A., Kuwabara, P.E. (2001) RNA-mediated interference as a tool for identifying drug targets. *Am. J. Pharmacogenomics* 1:45-53.

50. Peelle, B., Gururaja, T.L., Payan, D.G., Anderson, D.C. (2001) Characterization and use of green fluorescent proteins from *Renilla mulleri* and *Ptilosarcus guernyi* for the human cell display of functional peptides. *J. Protein Chem.* 20:507-519.

51. Pertl, U., Wodrich, H., Ruelmann, J.M., Gillies, S.D., Lode, H.N., Reisfeld, R.A. (2003) Immunotherapy with a posttranscriptionally modified DNA vaccine induces complete protection against metastatic neuroblastoma. *Blood* 101:649-654.

52. Phillips, M.I., ed. (1999a) Antisense Technology, Part A. Methods in Enzymology Vol. 313. Academic Press, Inc.

53. Phillips, M.I., ed. (1999b) Antisense Technology, Part B. Methods in Enzymology Vol. 314. Academic Press, Inc.

54. Pinkert, C.A., ed. (1994) Transgenic Animal Technology: A Laboratory Handbook. Academic Press.

55. Remington, J.P. (1985) Remington's Pharmaceutical Sciences. 17th ed. Mack Publishing Co.

56. Samal, B., Sun, Y.H., Stearns, G., Xie, C.S., Suggs, S., and McNiece, I. (1994). Cloning and characterization of the cDNA encoding a novel human pre-B-cell colony enhancing factor. Mol. Cell. Biol. 14(2): 1431 – 1437.

57. Sambrook, J., Russell, D.W., Sambrook, J. (1989) Molecular Cloning, A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press.

58. Schoen, F. J., (1994), "Robbins Pathologic Basis of Disease," W.B. Saunders Company, Philadelphia.

59. Stein, C.A., Kreig, A.M., eds. (1998) Applied Antisense Oligonucleotide Technology. Wiley-Liss.

60. Tang, D.C., DeVit, M., Johnston, S.A. (1992) Genetic immunization is a simple method for eliciting an immune response. *Nature* 356:152-154.

61. Wagner, R.W., Matteucci, M.D., Grant, D., Huang, T., Froehler, B.C. (1996) Potent and selective inhibition of gene expression by an antisense heptanucleotide. *Nat. Biotechnol.* 14:840-844.

62. Wagner, R.W., Matteucci, M.D., Lewis, J.G., Gutierrez, A.J., Moulds, C., Froehler, B.C. (1993) Antisense gene inhibition by oligonucleotides containing C-5 propyne pyrimidines. *Science* 260:1510-1513.

63. Xu, C.W., Mendelsohn, A.R., Brent, R. (1997) Cells that register logical relationships among proteins. *Proc. Natl. Acad. Sci. (USA)* 94:12,473-12,478.

64. Xu, Y., Piston, D.W., Johnson, C.H. (1999) A bioluminescence resonance energy transfer (BRET) system: Application to interacting circadian clock proteins. *Proc. Natl. Acad. Sci.* 96:151-156.

65. Yu, Z., Restifo, N.P. (2002) Cancer vaccines: progress reveals new complexities. *J. Clin. Invest.* 110:289-294.

66. Zallipsky, S. (1995) Functionalized poly(ethylene glycols) for preparation of biologically relevant conjugates. *Bioconjugate Chem.*, 6:150-165.

67. Zhu, J., Kahn, C.R. (1997) Analysis of a peptide hormone-receptor interaction in the yeast two-hybrid system *Proc. Natl. Acad. Sci.* 94:13,063-13,068.

68. Zavyalov et al., 1997, AP 105(3):161-186.

Claim

[00305] 1. A heterologous polypeptide comprising a secretory leader and a mature polypeptide, wherein the secretory leader is operably linked to an N-terminus of the mature polypeptide, wherein the secretory leader is not so linked to the mature polypeptide in nature, and wherein the secretory leader comprises a leader sequence of a secreted protein, and the secreted protein is selected from Table 1.

Abstract

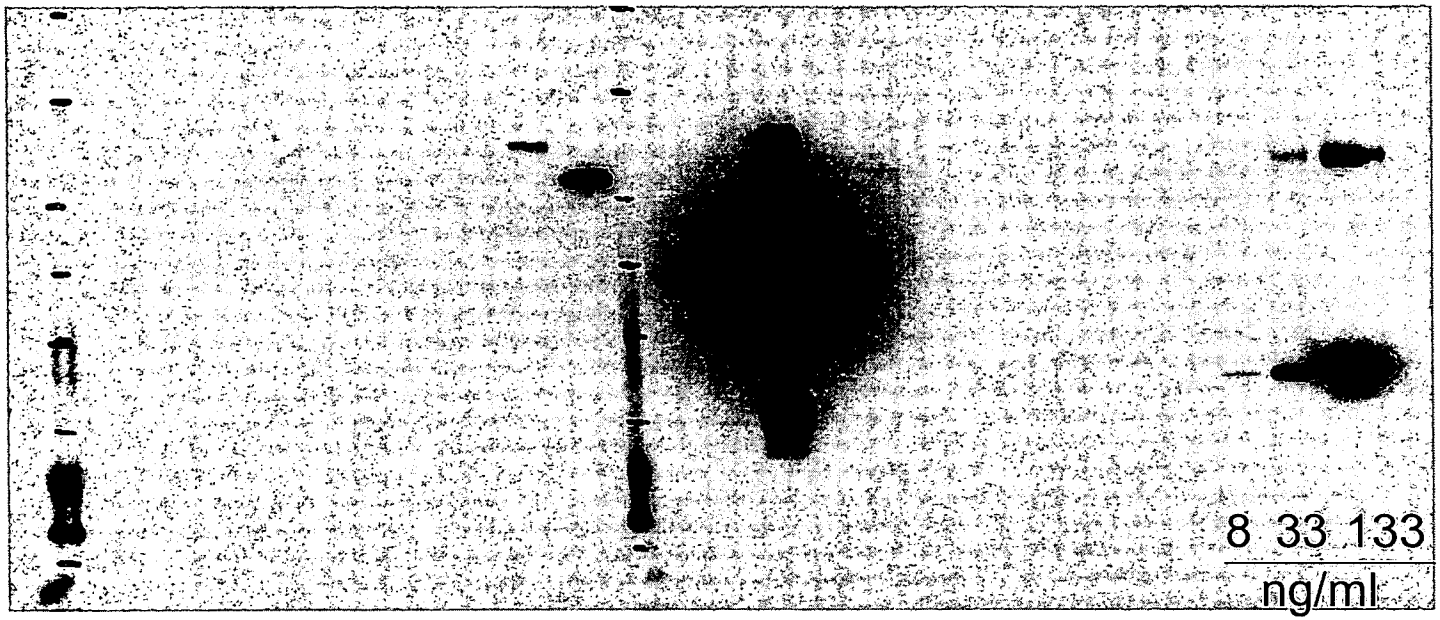
[00306] The present invention provides leader sequences that are useful for the production of heterologous secreted polypeptides, nucleic acid constructs that encode such leader sequences and heterologous secreted polynucleotides, vectors that contain such nucleic acid constructs, recombinant host cells that contain such nucleic acid constructs, vectors and polypeptides, and methods of making and using such secreted polypeptides with such heterologous leader sequences.

FIGURE 1 (CONT)

```
c.NP_001842_NM_001851  GEKGARGLDGEPGPQGLPGAPGDQGQRGPPGEAGPKGDRGAEGARGIPGLPGPKGDTGLP 540

b.CLN00517648_5pv1     ------------------------------------------------------------ 253
a.collagen_leader      ------------------------------------------------------------ 23
c.NP_001842_NM_001851  GVDGRDGIPGMPGTKGEPGKPGPPGDAGLQGLPGVPGIPGAKGVAGEKGSTGAPGKPGQM 600

b.CLN00517648_5pv1     ------------------------------------------------------------ 253
a.collagen_leader      ------------------------------------------------------------ 23
c.NP_001842_NM_001851  GNSGKPGQQGPPGEVGPRGPQGLPGSRGELGPVGSPGLPGKLGSLGSPGLPGLPGPPGLP 660

b.CLN00517648_5pv1     ------------------------------------------------------------ 253
a.collagen_leader      ------------------------------------------------------------ 23
c.NP_001842_NM_001851  GMKGDRGVVGEPGPKGEQGASGEEGEAGERGELGDIGLPGPKGSAGNPGEPGLRGPEGSR 720

b.CLN00517648_5pv1     ------------------------------------------------------------ 253
a.collagen_leader      ------------------------------------------------------------ 23
c.NP_001842_NM_001851  GLPGVEGPRGPPGPRGVQGEQGATGLPGVQGPPGRAPTDQHIKQVCMRVIQEHFAEMAAS 780

b.CLN00517648_5pv1     ------------------------------------------------------------ 253
a.collagen_leader      ------------------------------------------------------------ 23
c.NP_001842_NM_001851  LKRPDSGATGLPGRPGPPGPPGPPGENGFPGQMGIRGLPGIKGPPGALGLRGPKGDLGEK 840

b.CLN00517648_5pv1     ------------------------------------------------------------ 253
a.collagen_leader      ------------------------------------------------------------ 23
c.NP_001842_NM_001851  GERGPPGRGPNGLPGAIGLPGDPGPASYGKNGRDGERGPPGLAGIPGVPGPPGPPGLPGF 900

b.CLN00517648_5pv1     ---------------------- 253
a.collagen_leader      ---------------------- 23
c.NP_001842_NM_001851  CEPASCTMQAGQRAFNKGPDP 921
```

FIGURE 2

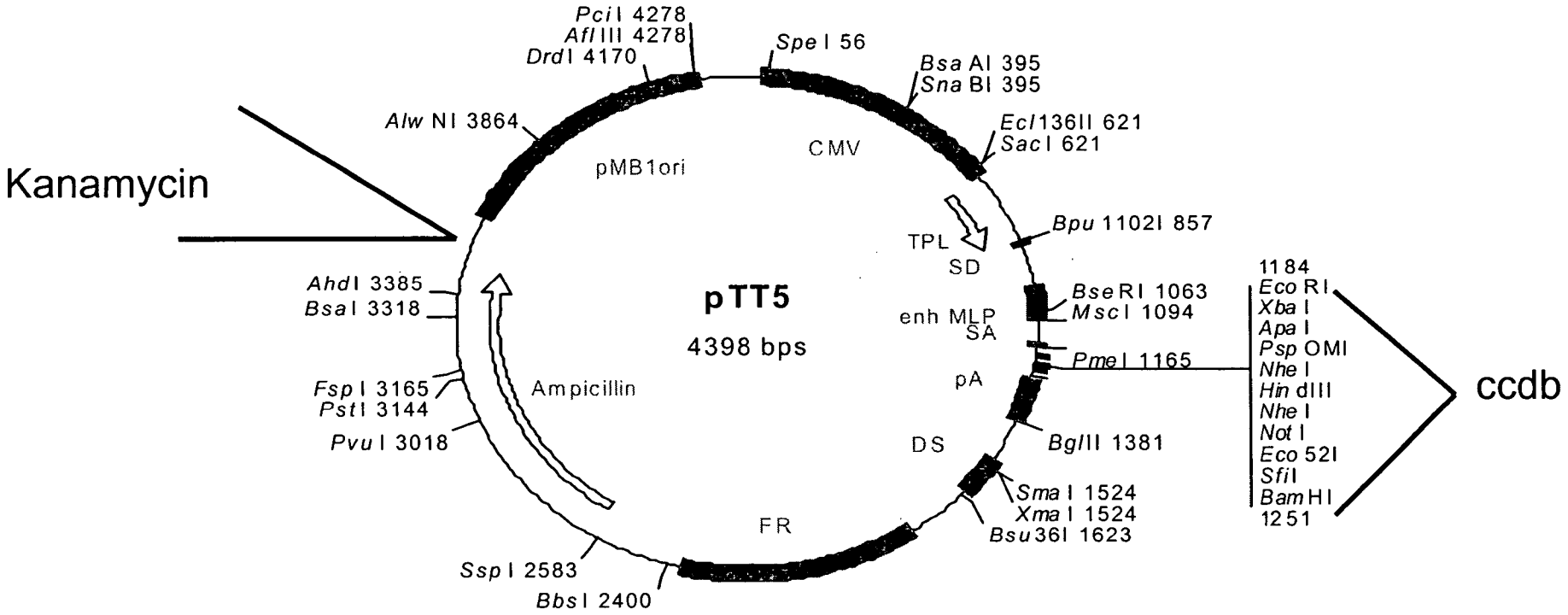
Experimental SP



CLN00517648 Expression in Media

FIGURE 3

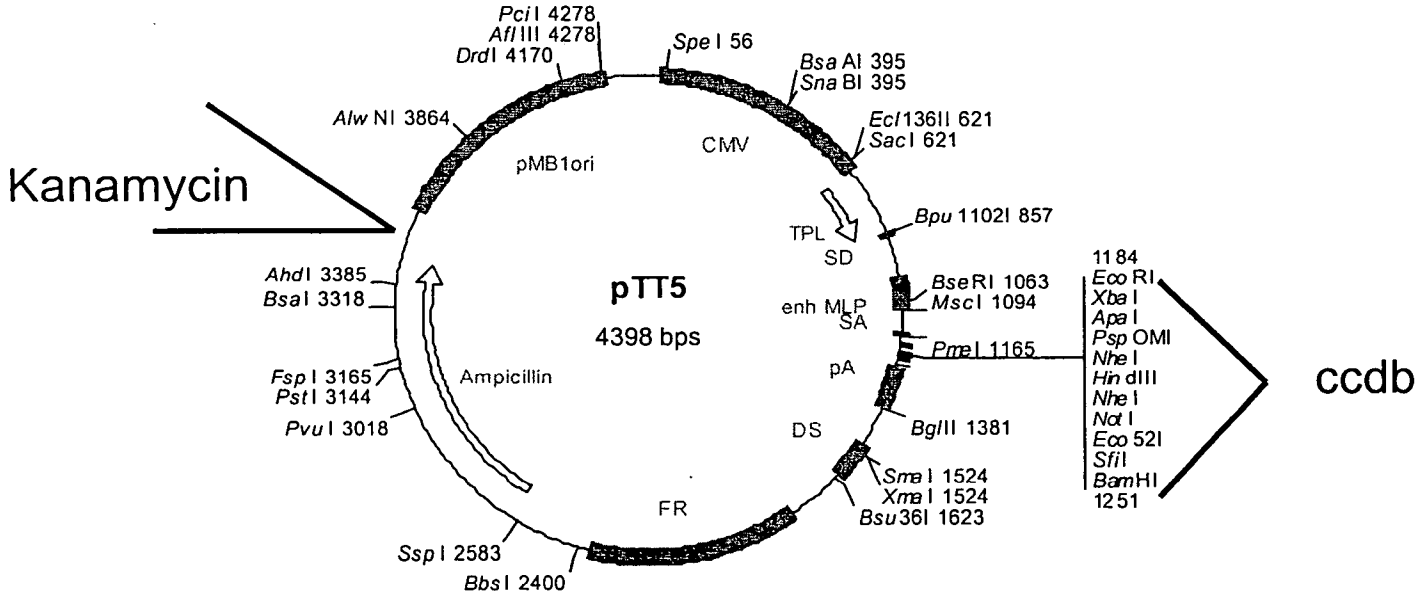
Vector for Type II cloning the Soluble Portion of the Type II Proteins: Back Bone Modifications



Yves Durocher's vector

FIGURE 4

Vector for Secreted Proteins with a CleavableTag



Vector A:

GAATTC----------GGATCCCTGGTTCCGCGTGGCTCAGGCTCATTCGAAGGTAAGCCTATCCC

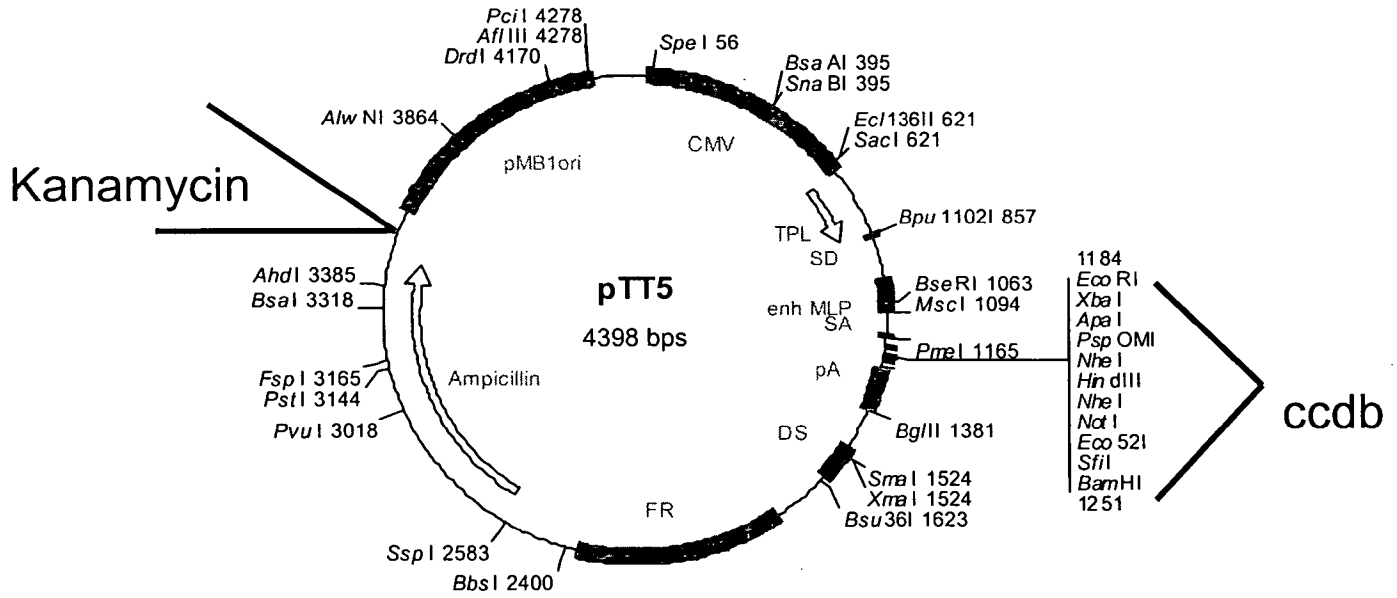
EcoR1 BamH1 Thrombin

TAACCCTCTCCTCGGTCTCGATTCTACGCGTACCGGTCATCATCACCATCACCATCACCATGGAGGACAGTGA

V5H8

FIGURE 5

Vectors for Producing Secreted Proteins with and w/o a CleavableTag



Vector B:

GCCGCCACCATGAAGACCTGCTGGAAAATTCCAGTTTTCTTCTTTGTGTGCAGTTTCCTGGAACCCTGGGCATCT
Kozak SP

GCAGAATTC---------GGATCCTTCGAAGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGCGTACCGGT
EcoR1 BamH1 V5H8

CATCATCACCATCACCATCACCATGGAGGACAGTGA

Vector C:

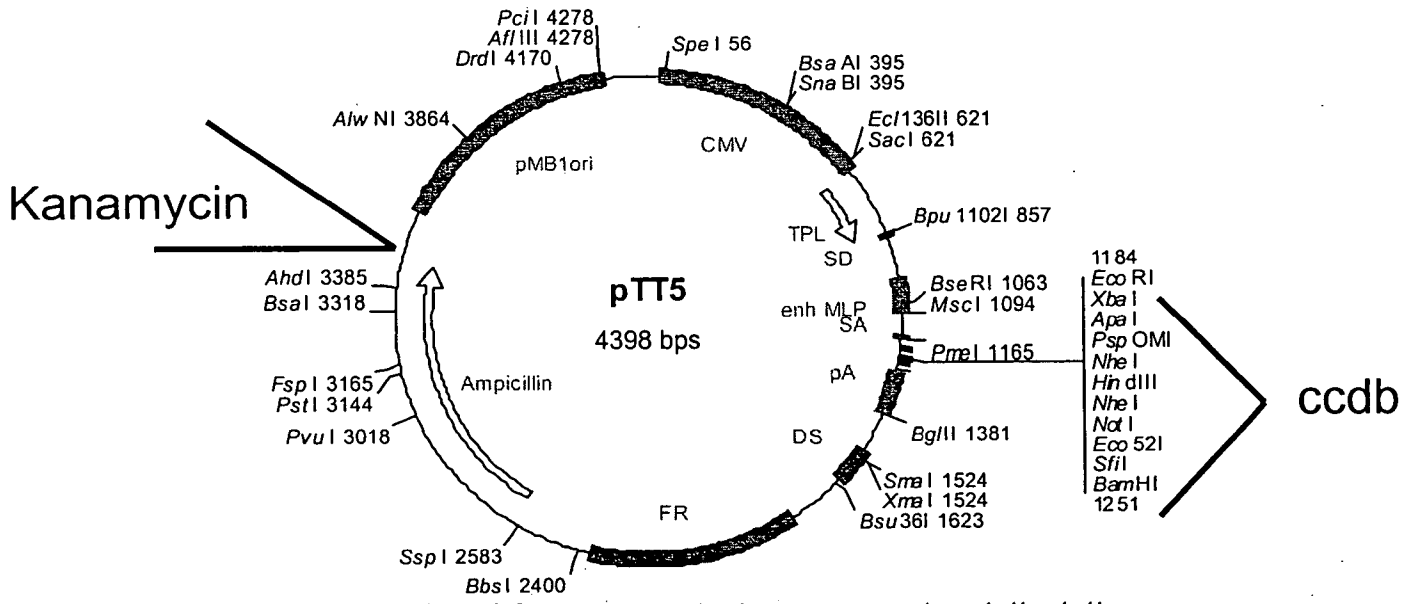
GCCGCCACCATGAAGACCTGCTGGAAAATTCCAGTTTTCTTCTTTGTGTGCAGTTTCCTGGAACCCTGGGCATCT
Kozak SP

GCAGAATTC---------GGATCCCTGGTTCCGCGTGGCTCAGGCTCATTCGAAGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTC
EcoR1 BamH1 Thrombin V5H8

GATTCTACGCGTACCGGTCATCATCACCATCACCATCACCATGGAGGACAGTGA

FIGURE 6

Vectors D and E



Vector D:

GAATTC----------GGATCCgcacctgaactcctgggggaccgtcagtcttcctcttccccccaaaacccaaggaca

EcoR1 BamH1 Fc

ccctcatgatctcccggacccctgaggtcacatgcgtggtggtggacgtgagccacgaagaccctgaggtcaagttcaactggtacgtggacggcgtggaggtgcataatgccaa
gacaaagccgcgggaggagcagtacaacagcacgtaccgtgtggtcagcgtcctcaccgtcctgcaccaggactggctgaatggcaaggagtacaagtgcaaggtctccaac
aaagccctcccagccccatcgagaaaaccatctccaaagccaaagggcagccccgagaaccacaggtgtacaccctgccccatcccgggatgagctgaccaagaacca
ggtcagcctgacctgcctggtcaaaggcttctatcccagcgacatcgccgtggagtgggagagcaatgggcagccggagaacaactacaagaccacgcctcccgtgctggact
ccgacggctccttcttcctctacagcaagctcaccgtggacaagagcaggtggcagcaggggaacgtcttctcatgctccgtgatgcatgaggctctgcacaaccactacacaca
gaagagcctctccctgtctccgggtaaatga

Vector E:

GCCGCCACCATGAAGACCTGCTGGAAAATTCCAGTTTTCTTCTTTGTGTGCAGTTTCCTGGAACCCTGGGCATCT

SP

GCAGAATTC--------- GGATCCgcacctgaactcctgggggaccgtcagtcttcctcttccccccaaaacccaaggacaccctcatgatctcccggacccc

EcoR1 BamH1 Fc

tgaggtcacatgcgtggtggtggacgtgagccacgaagaccctgaggtcaagttcaactggtacgtggacggcgtggaggtgcataatgccaagacaaagccgcgggaggagc
agtacaacagcacgtaccgtgtggtcagcgtcctcaccgtcctgcaccaggactggctgaatggcaaggagtacaagtgcaaggtctccaacaaagccctcccagccccatcga
gaaaaccatctccaaagccaaagggcagccccgagaaccacaggtgtacaccctgccccatcccgggatgagctgaccaagaaccaggtcagcctgacctgcctggtcaaa
ggcttctatcccagcgacatcgccgtggagtgggagagcaatgggcagccggagaacaactacaagaccacgcctcccgtgctggactccgacggctccttcttcctctacagcaa
gctcaccgtggacaagagcaggtggcagcaggggaacgtcttctcatgctccgtgatgcatgaggctctgcacaaccactacacacagaagagcctctccctgtctccgggtaaat
ga

FIGURE 7

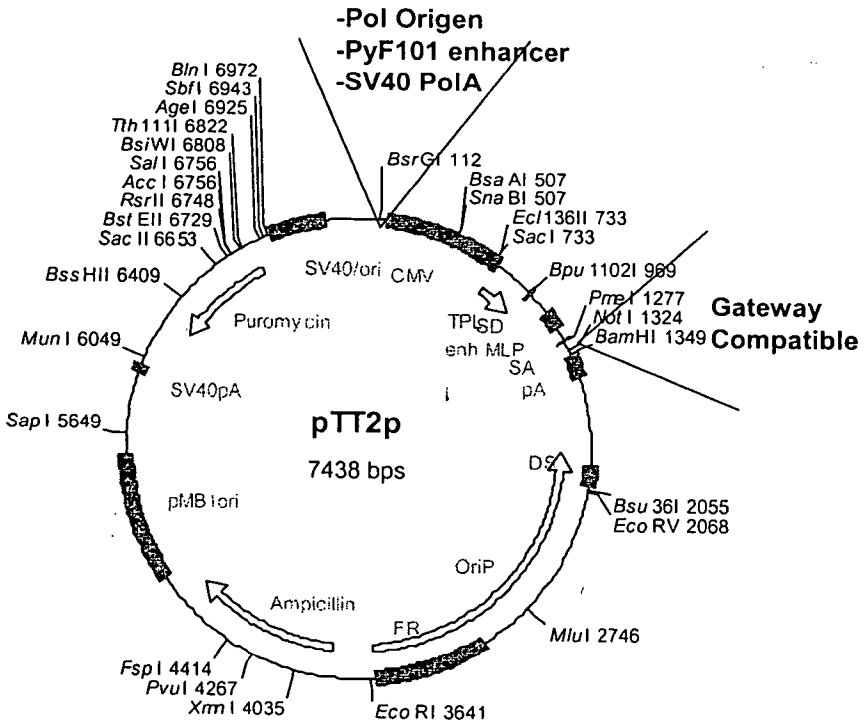
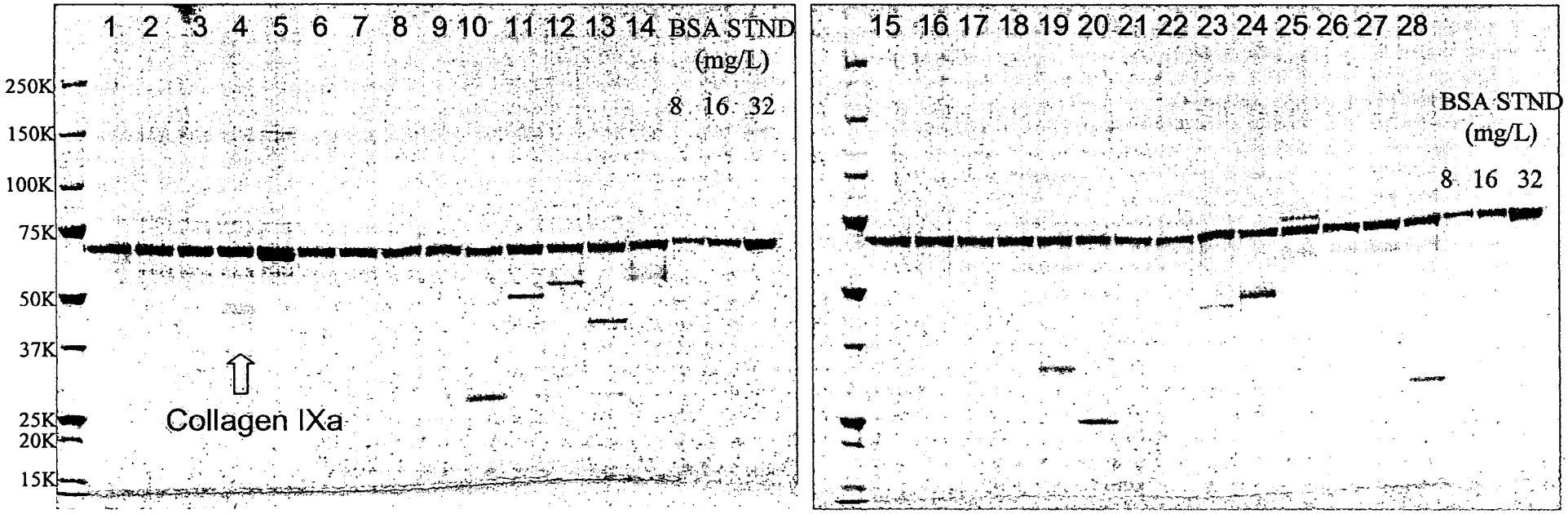
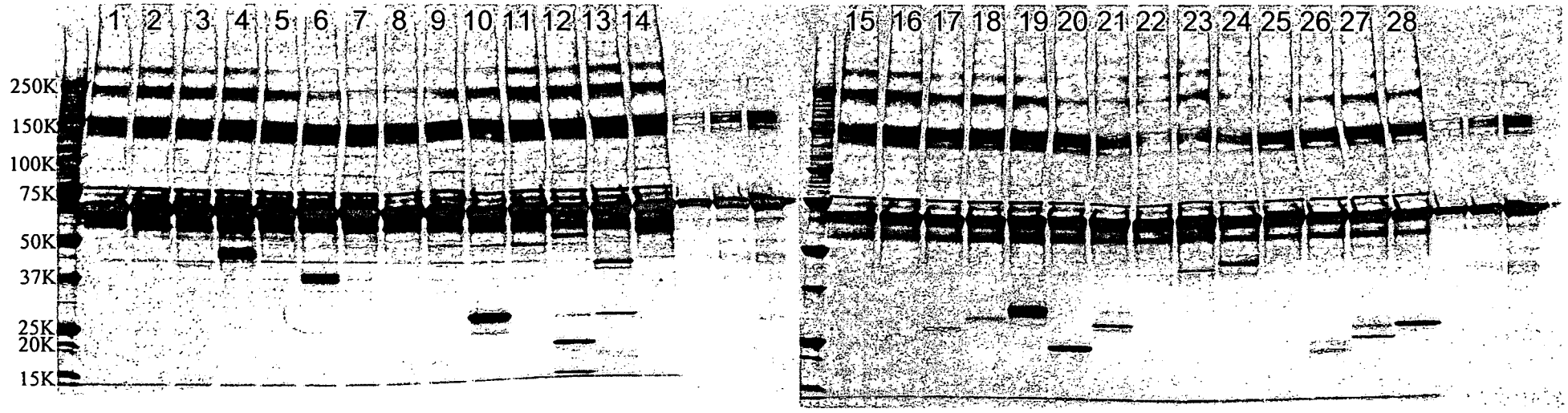

-Pol Origen
-PyF101 enhancer
-SV40 PolA
Bln I 6972
Sbf I 6943
Age I 6925
Tth111I 6822
BsiWI 6808
Sal I 6756
Acc I 6756
Rsr II 6748
Bst EII 6729
Sac II 6653
Bss HII 6409
Mun I 6049
Sap I 5649
BsrGI 112
Bsa AI 507
Sna BI 507
Ecl136II 733
Sac I 733
Bpu 1102I 989
Pme I 1277
Not I 1324
BamHI 1349
Gateway
Compatible
SV40/ori
CMV
TPL
SD
enh MLP
SA
pA
Puromycin
SV40pA
pTT2p
7438 bps
pMB1ori
DS
Bsu 36I 2055
Eco RV 2068
OriP
Ampicillin
FR
MluI 2746
Fsp I 4414
Pvu I 4267
Xmn I 4035
Eco RI 3641

FIGURE 8

SDS-PAGE Analysis of High Expressor Clone (48H Sup in CHO Soy)

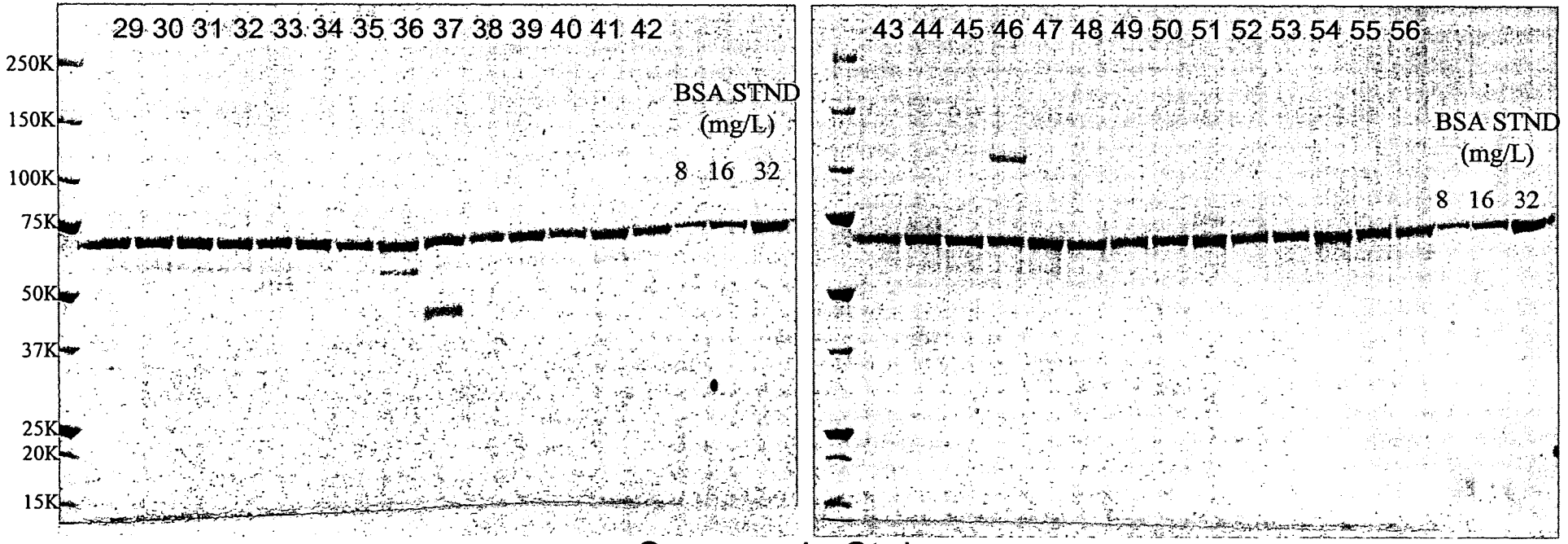


Coomassie Stain

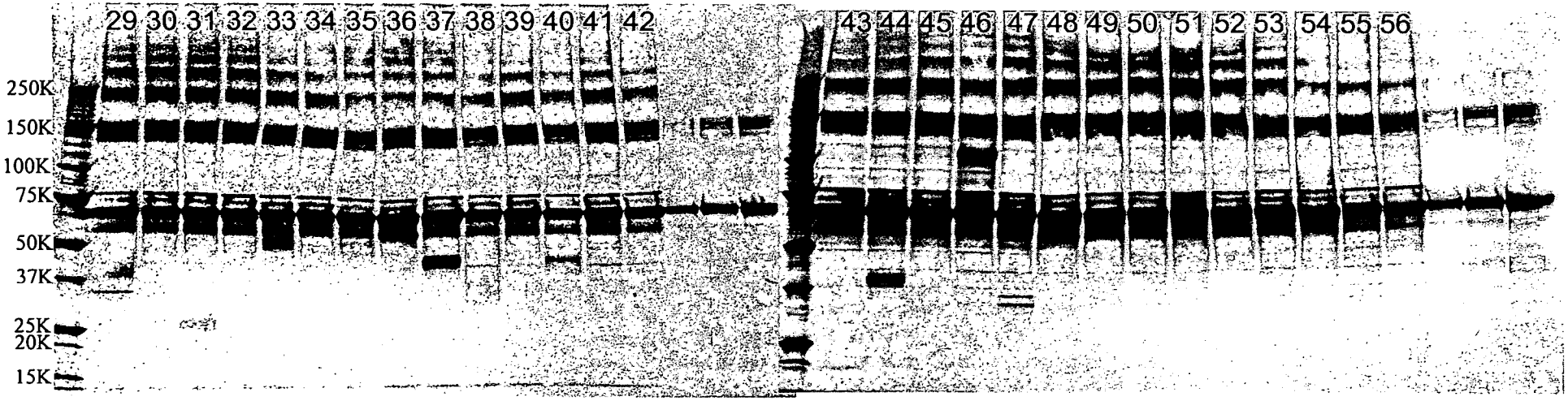


Silver Stain

FIGURE 9



Coomassie Stain



Silver Stain